System Requirements for SAS® Viya®

2021.2.4

This document might apply to additional versions of the software. Open this document in SAS Help Center and click on the version in the banner to see all available versions.

Virtual Infrastructure Requirements

Help with Cluster Setup

SAS Viya deployment requires experience with Kubernetes. However, SAS provides tools to help administrators create and configure a cluster that meets SAS Viya system requirements.

The SAS Viya Infrastructure as Code (IaC) projects contain scripts and configuration files that can automatically provision the infrastructure components that are required to deploy SAS Viya on Microsoft Azure, on Amazon Web Services (AWS), and on Google Cloud Platform (GCP). Each toolkit was developed to meet the same system requirements that are documented here.

Note: SAS Viya IaC tools are not available for a deployment on Red Hat OpenShift, and use of these tools does not support a deployment with multi-tenancy enabled at this time.

Some knowledge of Kubernetes and relevant third-party tools is still required. Use of the IaC tools means that some of the procedures in this guide do not apply or require modifications in order to deploy SAS Viya.

For more information, see these GitHub projects:

- SAS Viya 4 Infrastructure as Code (IaC) for Microsoft Azure
- SAS Viya 4 Infrastructure as Code (IaC) for AWS
- SAS Viya 4 Infrastructure as Code (IaC) for GCP

For most deployments, SAS recommends using the IaC tools to create the cluster and then using the SAS Deployment Operator to automate the deployment and future updates.

Required Permissions

Deploying the software and running tasks to operate SAS Viya servers require administrative access to the cluster and to the one or more namespaces that are used. For example, any user who issues kubectl commands to operate or check the status of SAS Viya servers will need this level of access. In the SAS Viya documentation, this level of access is referred to as *elevated Kubernetes permissions*.

For more information, see the Kubernetes documentation on role-based access control.

Host Machine and Cluster Requirements

SAS Viya runs in a Kubernetes cluster on multiple cloud platforms. At this time, only the cloud providers that are specified in Kubernetes Cluster Requirements are supported. Note that SAS does not support the on-premises Kubernetes services that are supplied by the public cloud providers.

SAS provides Limited Support for SAS Viya when it is deployed on a distribution of Kubernetes that is not listed here. See https://support.sas.com/en/technical-support/services-policies.html#k8s for the detailed support policy that applies to Kubernetes.

Virtual Private Cloud Considerations

SAS Viya supports multiple VPCs (or virtual networks) and multiple subnets per deployment. Although public IP addresses are required only for the load balancer and the ingress controller, SAS Viya pods and services consume many private IP addresses. When you plan your deployment, dedicate at least one VPC with subnets that are configured with sufficiently broad CIDR ranges to accommodate these private IP addresses.

Your Kubernetes service uses preferred methods for scaling up deployments with multiple IP addresses. For example, GKE supports VPC-native clusters, in which pod IP addresses are not dependent on static routes. For AKS, SAS has tested with a private CIDR and added a secondary CIDR to support larger deployments with multiple SAS Viya offerings. With EKS, consider setting up a very broad CIDR or a separate secondary IP address range in an additional subnet. For all platforms, consider using separate CIDR blocks for pods and for services, rather than placing pods and services in the same CIDR block.

A CNI plug-in can be a helpful addition for IP address management in a Kubernetes cluster.

Kubernetes Cluster Requirements

The Kubernetes cluster is where your SAS Viya software runs. Cluster requirements depend on the supported platform and are summarized in the following sections.

The optional SAS Viya Infrastructure as Code (IaC) projects create the Kubernetes cluster and set up the recommended node pools. They also install the ingress controller, cert-manager, and a load-balancer service. They can add these items to a cluster that they have set up, or to a cluster that you have set up as long as it complies with the requirements listed here.

Cluster Requirements for Microsoft Azure

The following table summarizes cluster requirements in a Microsoft Azure environment:

Table 1 *Cluster Requirements for Microsoft Azure*

Required Component	Detailed Requirements
Kubernetes	Microsoft Azure Kubernetes Service (AKS) 1.20.*x* - 1.22.*x*. **Note:** Be aware that some components, such as NGINX Ingress Controller and cert-manager, require upgrades to newer releases for use with Kubernetes 1.22.*x*. Check the appropriate third-party documentation for these compatibility requirements. In addition, refer to "A Note on Upgrading Kubernetes" for a workflow that avoids issues. **IMPORTANT** If you plan to upgrade Kubernetes from 1.21 to 1.22 and have an active SAS Viya deployment in place, be sure to stop the SAS Viya deployment before upgrading Kubernetes. Once the upgrade is complete, the SAS Viya deployment can be safely started.

Required Component	Detailed Requirements
A dedicated namespace per deployment	Sharing a namespace with other software deployments is not generally supported. However, a namespace can be shared with software that SAS recommends in order to support logging, monitoring, and other features. One method of uninstalling a SAS Viya deployment includes deleting the dedicated namespace. For more information, see "Uninstalling" in *SAS Viya: Deployment Guide*. You can determine what namespaces are defined in your cluster by running the following command: `kubectl get namespaces`
An ingress controller	NGINX Ingress Controller version 0.41.0 and later are supported. Istio Ingress and Istio Gateway are not supported at this time. Some additional requirements apply. For more information, see "Additional Configuration for Ingress and Load Balancers".
Kubernetes LoadBalancer services	A Kubernetes LoadBalancer service is created for the ingress controller when it is installed. It triggers the creation of an external IP address that is associated to an external load balancer, both of which are provisioned by the cloud provider. This IP address is used for all external HTTP/HTTPS connections to SAS Viya user interfaces and services. See the section on the LoadBalancer service in the Kubernetes documentation for more information. Additional Kubernetes load balancers are required to enable external user access to the CAS controller or to SAS/CONNECT. Some additional requirements apply. For more information, see "Additional Configuration for Ingress and Load Balancers".

Node pools	Cloud providers use different terminology to refer to collections of nodes. A minimum of three *dynamic node pools* is required:
	
	- One default node pool
	
	  Because SAS Viya software is not intended to run on these nodes, minimal resources are required.
	
	- Two user node pools
	
	  One of these node pools must be fully dedicated to the CAS server components. After the deployment has completed, if you find that SAS Viya software is landing on the default node pool in your configuration, reevaluate the size of the VMs in these user node pools. Sizing guidance is provided in "Resource Guidelines".
	
	SAS recommends creating six node pools, including the default node pool.
	
	Node pools that span multiple availability zones are not recommended and require additional configuration. For example, Azure managed disks (the default storage class for

Required Component	Detailed Requirements
	AKS) are not zone-redundant. In this case, SAS recommends that you follow the advice in this Microsoft document.
A certificate generator to enable TLS	By default, cert-manager is used if you have installed cert-manager in the cluster. However, you can use openssl instead. For more information, see "TLS Requirements".

Cluster Requirements for AWS

Most SAS Viya offerings support deployment with Amazon Elastic Kubernetes Service. Exceptions are noted in "Software Offerings" in *Getting Started with SAS Viya Operations*.

The following table summarizes cluster requirements in an AWS environment:

Table 2 *Cluster Requirements for AWS*

Required Component	Detailed Requirements
Kubernetes	Amazon Elastic Kubernetes Service (Amazon EKS) 1.20.*x* - 1.22.*x*. **Note:** Be aware that some components, such as NGINX Ingress Controller and cert-manager, require upgrades to newer releases for use with Kubernetes 1.22.*x*. Check the appropriate third-party documentation for these compatibility requirements. In addition, refer to "A Note on Upgrading Kubernetes" for a workflow that avoids issues.
A Kubernetes Metrics Server	SAS Viya was tested with version 0.4.1 of the Metrics Server. For more information, see https://github.com/kubernetes-sigs/metrics-server.
A dedicated namespace per deployment	Sharing a namespace with other software deployments is not generally supported. However, a namespace can be shared with software that SAS recommends in order to support logging, monitoring, and other features. One method of uninstalling a SAS Viya deployment includes deleting the dedicated namespace. For more information, see "Uninstalling" in *SAS Viya: Deployment Guide*. You can determine what namespaces are defined in your cluster by running the following command: `kubectl get namespaces`
An NGINX ingress controller	NGINX Ingress Controller version 0.41.0 and later are supported. Istio Ingress and Istio Gateway are not supported at this time. Some additional requirements apply. For more information, see "Additional Configuration for Ingress and Load Balancers".

Required Component	Detailed Requirements
Kubernetes LoadBalancer services	A Kubernetes LoadBalancer service is created for the ingress controller when it is installed. It triggers the creation of an external IP address that is associated to an external load balancer, both of which are provisioned by the cloud provider. This IP address is used for all external HTTP/HTTPS connections to SAS Viya user interfaces and services. See the section on the LoadBalancer service in the Kubernetes documentation for more information. Additional Kubernetes load balancers are required to enable external user access to the CAS controller or to SAS/CONNECT. Some additional requirements apply. For more information, see "Additional Configuration for Ingress and Load Balancers".
Managed node groups	Cloud providers use different terminology to refer to collections of nodes. A minimum of three *dynamic node pools* is required: ■ One default node pool Because SAS pods are not intended to run on these nodes, minimal resources are required. ■ Two additional node pools for SAS pods One of these node pools must be fully dedicated to the CAS server components. After the deployment has completed, if you find that SAS Viya pods are landing on the default node pool in your configuration, reevaluate the size of the machines in these user node pools. Sizing guidance is provided in "Resource Guidelines". SAS recommends creating six node pools, including the default node pool. Node pools that span multiple availability zones (AZs) are not recommended.
A certificate generator to enable TLS	By default, cert-manager is used if you have installed cert-manager in the cluster. However, you can use the OpenSSL-based certificate generator instead. For more information, see "TLS Requirements".

Cluster Requirements for GCP

Most SAS Viya offerings support deployment with Google Kubernetes Engine. Exceptions are noted in "Software Offerings" in *Getting Started with SAS Viya Operations*.

The following table summarizes cluster requirements in a GCP environment:

Table 3 *Cluster Requirements for Google Cloud Platform*

Required Component	Detailed Requirements
Kubernetes	Google Kubernetes Engine (GKE) 1.20.*x* - 1.22.*x*. **Note:** Be aware that some components, such as NGINX Ingress Controller and cert-manager, require upgrades to newer releases for use with Kubernetes 1.22.*x*. Check the appropriate third-party documentation for these compatibility requirements. In addition, refer to "A Note on Upgrading Kubernetes" for a workflow that avoids issues.
A dedicated namespace per deployment	Sharing a namespace with other software deployments is not generally supported. However, a namespace can be shared with software that SAS recommends in order to support logging, monitoring, and other features. One method of uninstalling a SAS Viya deployment includes deleting the dedicated namespace. For more information, see "Uninstalling" in *SAS Viya: Deployment Guide*. You can determine what namespaces are defined in your cluster by running the following command: `kubectl get namespaces`
An NGINX ingress controller	NGINX Ingress Controller version 0.41.0 and later are supported. Istio Ingress and Istio Gateway are not supported at this time. Some additional requirements apply. For more information, see "Additional Configuration for Ingress and Load Balancers".
Kubernetes LoadBalancer services	A Kubernetes LoadBalancer service is created for the ingress controller when it is installed. It triggers the creation of an external IP address that is associated to an external load balancer, both of which are provisioned by the cloud provider. This IP address is used for all external HTTP/HTTPS connections to SAS Viya user interfaces and services. See the section on the LoadBalancer service in the Kubernetes documentation for more information. Additional Kubernetes load balancers are required to enable external user access to the CAS controller or to SAS/CONNECT. Some additional requirements apply. For more information, see "Additional Configuration for Ingress and Load Balancers".
Node pools	Cloud providers use different terminology to refer to collections of nodes. A minimum of three *dynamic node pools* is required: ■ One default node pool Because SAS pods are not intended to run on these nodes, minimal resources are required. ■ Two additional node pools for SAS pods

Required Component	Detailed Requirements
	One of these node pools must be fully dedicated to the CAS server components. After the deployment has completed, if you find that SAS Viya pods are landing on the default node pool in your configuration, reevaluate the size of the VMs in these user node pools. Sizing guidance is provided in "Resource Guidelines". SAS recommends creating six node pools, including the default node pool. Node pools that span multiple zones are not recommended.
A certificate generator to enable TLS	By default, cert-manager is used if you have installed cert-manager in the cluster. However, you can use the OpenSSL-based certificate generator instead. For more information, see "TLS Requirements".

Cluster Requirements for Red Hat OpenShift

Most SAS Viya offerings support deployment with Red Hat OpenShift. Exceptions are noted in "Software Offerings" in *Getting Started with SAS Viya Operations*.

The following table summarizes cluster requirements in a Red Hat OpenShift environment:

Table 4 *Cluster Requirements for OpenShift*

Required Component	Detailed Requirements
Kubernetes	Red Hat OpenShift Container Platform (OCP) 4.7.*x* on VMware vSphere 7.0.1. SAS has tested only with a user-provisioned infrastructure installation.
Compliant machines	Red Hat Enterprise CoreOS (RHCOS) is the only operating system that Red Hat supports for OCP control plane nodes. Red Hat enables you to use Red Hat Enterprise Linux on the worker nodes, but SAS has tested only with RHCOS.
A dedicated namespace per deployment	Sharing a namespace with other software deployments is not generally supported. However, a namespace can be shared with software that SAS recommends in order to support logging, monitoring, and other features. One method of uninstalling a SAS Viya deployment includes deleting the dedicated namespace. For more information, see "Uninstalling" in *SAS Viya: Deployment Guide*. You can determine what namespaces are defined in your cluster by running the following command: `kubectl get namespaces`
Cluster ingress	Only the OpenShift Ingress Operator is supported.

Required Component	Detailed Requirements
	You must modify the kustomization.yaml file to enable routes. For more information, see "Create the File" in *SAS Viya: Deployment Guide*. Verify that the requirements in "Additional Configuration for Ingress and Load Balancers" have also been met.
Kubernetes LoadBalancer services	Usage varies based on the underlying infrastructure. See the endpointPublishingStrategy configuration parameter section of the OpenShift Ingress Operator documentation for information about whether a LoadBalancer service is used by your cluster infrastructure.
Node labels	One or more nodes should be fully dedicated to (that is, labeled and tainted for) the CAS server. This recommendation depends on whether your deployment uses the CAS server. SAS strongly recommends labeling at least one node for compute workloads. It is helpful to understand the workload placement strategy that is described in "Plan the Workload Placement" in *SAS Viya: Deployment Guide*.
A certificate generator to enable TLS	By default, cert-manager is used if you have installed cert-manager in the cluster. However, you can use the OpenSSL-based certificate generator instead. For more information, see "TLS Requirements".
cert-utils-operator	This operator from Red Hat is required to manage certificates for TLS support and create keystores. For more information, see https://github.com/redhat-cop/cert-utils-operator/blob/master/README.md.

A Note on Upgrading Kubernetes

If you are planning to upgrade your cluster to Kubernetes 1.22, SAS recommends taking the following steps in order to avoid issues:

1. First, update the SAS Viya software. SAS Viya must be at the 2021.2.4 (February 2022) stable release or later.
2. Verify that NGINX Ingress Controller and cert-manager (if you are using it for TLS certificate management) have been upgraded to versions that are compatible with Kubernetes 1.22.
3. Upgrade the Kubernetes cluster to version 1.22.

Kubernetes Client Machine Requirements

The deployment requires a machine from which the Kubernetes command-line interface, kubectl, manages a Kubernetes cluster. This machine can be running Linux, Windows, or macOS. It requires the following client software:

- kubectl

  The kubectl version that you use on the client machine can be only one minor version later or earlier than the version of Kubernetes (kube-apiserver) that is used in the cluster. For more information, see the Kubernetes version skew policy.

  Run the following command to verify the version of kubectl on the client machine:

  ```
  kubectl version --short
  ```

  After kubectl is installed, note the location of the Kubernetes configuration file for use during the deployment process. The default path on Linux is **~/.kube/config**.

- Kustomize 3.7.0

  Kustomize is a client tool that is used to generate Kubernetes manifest files. Each SAS Viya release and cadence is optimized for and tested with a single version of Kustomize.

  Run the following command to verify the version of Kustomize on the client machine:

  ```
  kustomize version --short
  ```

  Download Kustomize here: https://github.com/kubernetes-sigs/kustomize/releases/tag/kustomize%2Fv3.7.0.

- (For users of the orchestration image only) A recent version of Docker or another container runtime

  The sas-orchestration image includes a supported version of kubectl and Kustomize to help you deploy and manage SAS Viya. If you do not use the orchestration image for your deployment, you can also get these versions by following the instructions in the kubernetes-tools README. After you have downloaded the deployment assets for your software order, you can find it in **$deploy/sas-bases/examples/kubernetes-tools/README.md** or in HTML format in **sas-bases/docs/using_kubernetes_tools_from_the_sas-orchestration_image.htm**.

Node Requirements

The nodes in your Kubernetes cluster have the following requirements:

- A Linux operating system
- A 64-bit x86_64 chipset
- A recent version of Docker or another container runtime
- Kubernetes 1.20.*x* - 1.22.*x* installed on each node.

  To check your version of Kubernetes, run the following command:

  ```
  kubectl version
  ```

- One of the following requirements for Open Distro for Elasticsearch:
  - A cluster-wide setting to enable privileged containers
  - Increased virtual memory settings on the nodes that host stateful workloads

  For more information about these requirements, see “Open Distro for Elasticsearch Requirements”.

Additional machine requirements are described in “Resource Guidelines”.

Additional Configuration for Ingress and Load Balancers

The following additional settings are required for your ingress controller:

- A static public IP address, which is created automatically when the Kubernetes LoadBalancer service is created for the ingress controller.
- The public IP address or an equivalent address for the ingress controller must be registered with your DNS provider as an external endpoint.

  For Microsoft Azure and for GCP, an `A` record must point to the IP address.

  For AWS, a `CNAME` record must point to the fully qualified domain name (FQDN). AWS assigns `A` records (DNS FQDNs) to load balancers. For AWS, you should create the desired DNS name for each load balancer as a `CNAME` record and configure it with the FQDN.
- An external URL that is configured for the Ingress object and that is reachable from pods that run inside the cluster.

  Firewall rules that allow these outbound connections are required.
- Routes to the ingress that do not pass through intervening firewalls that perform network address translation (NAT). NAT placement before the ingress causes the SAS Viya audit service to report incorrect remote addresses.

  In addition, any NodePort or LoadBalancer services that the cluster uses for ingress must have their externalTrafficPolicy set to `Local`. As a result, an ingress pod must be deployed to the node that is receiving connection requests. Be aware of these requirements if you are using kube-proxy or an ingress operator to handle connections to cluster nodes.
- A reasonable time-out setting on the ingress controller.

  The default time-out value might be too low. SAS recommends setting a 300-second time-out in most environments.
- Adequate buffer sizes for request and response headers.

  SAS Viya passes OAuth tokens in the request and response headers. The size of each token varies, depending on the number of group memberships that are associated with each user account. For NGINX, the default configuration uses 32 KB request headers and 16 KB response headers.

  For OpenShift, the default limit is 24 KB for both request and response headers.
- If a reverse proxy server is active between the network and the ingress controller, NGINX configuration changes are required. For more information, see “Additional Configuration for Proxy Environments”.

An ingress controller is not sufficient to enable user access to the CAS controller binary port or to the SAS/CONNECT port from outside the cluster. The connections are not HTTP connections. Therefore, you must define a Kubernetes service of type *LoadBalancer* in order to make the non-HTTP ports externally accessible.

The following settings are required for your additional Kubernetes LoadBalancer services:

- A static public IP address.
- Static DNS names and ports.
- (For AWS only) A reasonable time-out value.

  The default time-out value might be too low. For example, the default time-out for AWS load balancers, 60 seconds, is too low. SAS recommends setting a 300-second time-out in most environments.

This setting can be changed for all of your CAS load balancers by specifying the `service.beta.kubernetes.io/aws-load-balancer-connection-idle-timeout` annotation in the metadata of the CASDeployment serviceTemplate. An example of the YAML for this modification is included in **`sas-bases/examples/cas/configure/cas-enable-external-services.yaml`**.

In some situations, users might connect directly to a node port or LoadBalancer service from an external IP address or host name, bypassing the ingress controller. Such cases might involve the use of SAS/CONNECT or the CAS programming interfaces, for example. To support TLS, additional configuration is required. For more information, see "TLS Requirements" on page 49.

Additional Configuration for Proxy Environments

An NGINX Ingress Controller requires additional configuration in environments where an application gateway or a reverse proxy server is set up in front of the cluster ingress.

A variable in the SAS Viya deployment manifest, `SAS_SERVICES_URL`, specifies the host name and port of the ingress controller for use by the compute server component. Similarly, to enable pods to reach the ingress controller, the value for the `{{ NAME-OF-INGRESS-HOST }}` parameter must match the value of the `INGRESS_HOST` variable. All these values should represent the "front door" to your cluster. Therefore, you can specify a load balancer or application gateway as the value for `SAS_SERVICES_URL` and `INGRESS_HOST`. This configuration is also appropriate for external reverse proxies. The load balancer can be for the ingress controller or external. For more information, see "Initial kustomization.yaml File" in *SAS Viya: Deployment Guide*.

If you are using a reverse proxy server with OpenID Connect and SCIM for authentication, the NGINX configuration setting `use-forwarded-headers` must be changed from the default "false" to "true". This change is required to enable an NGINX Ingress Controller to pass incoming X-Forwarded-* headers from the reverse proxy to SAS Viya services. For more information, see "SCIM Requirements".

Requirements for a Multi-Tenant Environment

Requirements to support multi-tenancy are included in various locations throughout this guide. If you are deploying SAS Viya for multiple tenants, be sure to meet the requirements that are specified in the following sections:

- [ ] CAS server. Each tenant requires a dedicated CAS server. See "CAS Server Resources" on page 29.

  The script that creates a CAS server for each tenant, create-cas-server.sh, requires Bash version 4 or later.

- [ ] SAS Infrastructure Data Server. Both the internal and external PostgreSQL options are supported for multi-tenancy. Some additional requirements apply. See "PostgreSQL Requirements for a Multi-Tenant Deployment" on page 66.

- [ ] TLS certificates. See "TLS Requirements" on page 49.

- [ ] DNS configuration. See "DNS Requirements for Multi-Tenancy" on page 50.

- [ ] User accounts in your LDAP or SCIM identity provider. Additional configuration is required. See "Additional LDAP Requirements for Multi-Tenancy" on page 52 or "Additional SCIM Requirements for Multi-Tenancy" on page 53.

Multi-tenancy is not supported in every customer environment. For more information, see "Limitations to Multi-Tenancy Support" on page 71.

Hardware and Resource Requirements

Storage Requirements

Use the information in this section to estimate the sizes of storage devices for your deployment.

Storage Overview

When selecting a storage option, keep in mind that your assortment of SAS product offerings, the number of users, and the amount of data that is processed all affect sizing requirements for performance. The performance of shared storage options is highly vendor-dependent.

SAS Viya requires both ephemeral storage and persistent storage volumes. Some SAS Viya components require shared storage. Network-based storage can perform better than locally mounted storage if bandwidth is adequate. However, SAS recommends provisioning the CAS server and compute nodes with both high-performing local storage and shared storage.

For more information about CAS storage requirements, see “Persistent Storage Volumes, PersistentVolumeClaims, and Storage Classes” on page 15.

The SAS Programming Run-Time Environment, which includes SAS Compute Server, SAS/CONNECT server, and SAS Batch Server, produces workloads in the compute workload class. These components must be able to create temporary files and data sets in a storage volume. For more information, see “SAS Programming Runtime Environment Requirements” on page 31.

SAS Viya Monitoring for Kubernetes is an optional solution for monitoring SAS Viya deployments. If deployed, it requires additional storage for storing log messages and performance metrics collected from the SAS Viya deployment. The amount of storage required by this monitoring solution is heavily dependent on your retention policies. For more information, see the SAS Viya Monitoring for Kubernetes project in GitHub: https://github.com/sassoftware/viya4-monitoring-kubernetes.

Instructions for modifying storage settings are provided in README files for several SAS Viya components. For example, the file titled “Configuration Settings for CAS” explains how to change the storage size for CAS PersistentVolumeClaims (PVCs), how to modify the resource allocation for ephemeral storage, and how to change the accessModes on the CAS permstore and data PVCs. After you have downloaded the deployment assets for your software order, the README files are located at `$deploy/sas-bases/examples/` (for Markdown format) or at `$deploy/sas-bases/docs/` (for HTML format). To modify storage for any component, first consult its README file.

I/O Throughput and Performance Considerations

The peak I/O throughput requirements of your SAS Viya deployment might exceed the capabilities of your storage configuration. SAS generally recommends sequential I/O bandwidth of 90-120 MB per second, per physical CPU core for both persistent and ephemeral storage. For best performance, select VM instance types with the highest available I/O throughput levels. On Microsoft Azure, Premium storage is required in order to achieve these I/O throughput levels.

Although it is high-performing, NVMe offers only ReadWriteOnce (RWO) storage that does not persist after a pod restarts. ReadWriteMany (RWX) storage is required for multiple SAS Viya components.

You should consult your cloud provider's storage and compute instance documentation to ensure that the storage environment can provide the level of I/O that is required. SAS Technical Support found that many performance issues reported by SAS Viya customers can be directly attributed to insufficient levels of I/O throughput.

In general, SSDs perform better than HDDs.

Disk Space

Most SAS Viya software uses the same base container and Docker layers in order to optimize disk usage. The minimum amount of disk space that is required for the installation and for logging is 48 GB. Therefore, the minimum combined capacity of the Kubernetes worker nodes should be 48 GB.

The number of containers and the size of the images depend on the products in the software order. These images require approximately 30-60 GB of disk space on the host on which you are downloading SAS images and also in the destination registry.

If you are using the optional SAS Mirror Manager, the same sizing guidelines apply to the machine where you run it. After the first run, SAS Mirror Manager creates a local copy of the container images to make the process of mirroring images as quick as possible.

The topics for individual platforms, beginning with Virtual Machine Recommendations for Microsoft Azure, provide guidelines for VM sizing, including recommended disk size.

Persistent Storage Volumes, PersistentVolumeClaims, and Storage Classes

Persistent storage is required by multiple SAS Viya components. Verify that at least one ReadWriteMany (RWX) StorageClass has been defined and set as your default. Run the following command to verify that a default StorageClass has been defined:

```
kubectl get storageclass
```

Make a note of the names of all storage classes in case you later need to modify storage settings. You must also update the base kustomization.yaml to specify the PVCs that are associated with the RWX StorageClass. For more information, see Specify PersistentVolumeClaims to Use ReadWriteMany StorageClass.

Many storage classes have a ReclaimPolicy that is set to **Delete** by default. If you delete a namespace that includes such storage classes, the PVCs in that namespace are deleted. If the ReclaimPolicy is Delete, the corresponding persistent volumes (PVs) are also deleted, resulting in data loss. For information about changing the ReclaimPolicy in Kubernetes, see Change the Reclaim Policy of a PersistentVolume.

Multiple PVCs are configured automatically during the deployment. Here are the default settings for all deployments. Note that the default settings correspond to the minimum recommended sizes:

- cacheserver pod — SAS Cache Server stores longer-term data, such as configuration settings, and serves it to SAS Viya services. It requires a storage volume. It is deployed with high availability (HA) by default, and two replicas, each with a 2 Gi PVC; accessMode: ReadWriteOnce (RWO).
- Consul — Supports SAS Configuration Server. It is deployed with high availability (HA) by default, and three replicas, each with a 1 Gi PVC; accessMode: RWO.

The Consul pod also requires a mount for the Consul data directory. The mount is an empty directory that points to `/consul/data`. It is configured automatically by the deployment process.

- RabbitMQ — Supports SAS Message Broker. It is deployed with HA by default and three replicas, each with a 2 Gi PVC; accessMode: RWO.
- PostgreSQL — Supports SAS Infrastructure Data Server. The requirements depend on whether you deploy the default (internal) PostgreSQL instance or you supply your own PostgreSQL server (external). Requirements are summarized in the following table:

Table 5 *SAS Infrastructure Data Server Storage Requirements*

PostgreSQL Deployment	accessMode	Default Size
Internal (the default deployment)	RWO	Deploys with HA by default: ■ Three nodes One primary and two replicant data nodes. ■ PVC of 128 GB per node
External (you are responsible for this server)	Not applicable	A minimum of one volume, 128 GB of space.

For more information, see “PostgreSQL Deployment Options”.

- CAS server — Requirements depend on whether you deploy the CAS server as SMP or MPP. Two PVCs are required in either case:
  - cas-default-permstore — Required in order to store caslib management privileges
  - cas-default-data — The default location for new caslibs; equivalent to the SAS Viya CASDATADIR setting

  Their requirements are summarized in the following table:

Table 6 *CAS Server Persistent Storage Requirements*

CAS Server	accessMode	Default Size
SMP CAS server	Either RWO or RWX. Set to RWX by default.	cas-default-permstore: 100 Mi cas-default-data: 8 Gi
MPP CAS server	RWX RWX	cas-default-permstore: 100 Mi cas-default-data: 8 Gi

RWX accessMode is required for any backup controllers for CAS. To change the accessMode for either cas-default-data or cas-default-permstore, perform the steps that are described in Change accessMode.

Every CAS pod also requires a mount for the CAS disk cache. The mount is configured automatically by the deployment process. CAS server performance is partially dependent on the storage that you select. For more information, see “CAS Server Resources” on page 29.

You can modify most default resources that are provided for your CAS server. However, the mountPath that is defined for cas-default-permstore and cas-default-data cannot be modified. The associated mount points are `/cas/permstore` and `/cas/data`, respectively. To define mounts that are added to CAS overlays, you can add patchTransformers to your manifests.

- Open Distro for Elasticsearch — Supports the search feature. Creates one PVC. Default size: 128 Gi; accessMode: RWO. For more information, see "Open Distro for Elasticsearch Requirements".
- Backup and restore operations — Require two PVCs, described as follows:
  - sas-common-backup-data — Stores a backup file of settings for SAS Infrastructure Data Server and SAS Configuration Server. Default size: 25 GB; accessMode: RWX.
  - sas-cas-backup-data — Stores a backup of CAS server data and the CAS permstore. Default size: 8 GB; accessMode: RWX.

> **IMPORTANT** These default settings are the minimum recommended sizes. Verify that the ReclaimPolicy for these PVCs is set to **Retain**.

Persistent Volumes for Applications

Some individual SAS Viya products also require persistent storage. If your software order included these applications, you must set up additional volumes with the required settings:

- SAS Common Planning Service — Requires two volumes:
  - sas-planning-retail — Stores data for offerings that include SAS Common Planning Service. Default size: 100Gi; accessMode: RWX.
  - sas-planning-retail-backup-data — Stores backup files for SAS Common Planning Service. Default size: 100Gi; accessMode: RWX.
- SAS Data Quality — Requires a volume for SAS Quality Knowledge Base data. Default size: 8 Gi; accessMode: RWX.
- SAS Micro Analytic Service — Requires storage volumes if the optional features, ASTORES or archives, are configured:
  - Volume for ASTORES — Default size: 30 GB; accessMode: RWX. Depending on model complexity and the number of models, more space might be required.

    To resize the PVC for ASTORES, update the overlay with the new size and run it. If models have already been loaded, SAS recommends setting the StorageClass attribute `allowVolumeExpansion` to true. If the provider does not support the StorageClass, resize the volume and then republish all ASTORE models.
  - Volume for archive logs — If the archive feature is enabled in SAS Micro Analytic Service, it stores input and output transaction logs in a persistent volume. Default size: 30 GB; accessMode: RWX.

    This volume grows continuously. SAS recommends that you create regular backups of this data and monitor the available space.
- SAS Studio — Requires persistent storage if users will take advantage of Git integration features.

  To enable Git integration, SAS Studio users require access to a shared file system that is accessible to the pod where the compute server is running (the pod where the compute workload has been placed). Be aware that storage on this pod is temporary unless you configure persistent storage. Configure persistent storage for the compute server if you plan to use the Git integration

features. For more information, see "Creating Persistent File Storage" in *SAS Studio: Administrator's Guide*.

Work that is associated with other SAS Studio tasks is stored in SAS Content, which does not require a persistent volume.

- SAS Configurator for Open Source — Requires persistent storage for Python builds. Default size: 20 GB; accessMode: RWX. The required size depends on the number of Python profiles that you configure, so more space might be required.

Additional PVC Requirements for Microsoft Azure

In a Microsoft Azure environment, any RWO persistent storage should be configured on VM instance types whose names include the "S" suffix. This suffix indicates that the VM instances have support for Premium (SSD) Storage.

By default, PVCs are owned by the root user. In Microsoft Azure environments, you must define a StorageClass in which you explicitly set mountOptions for each PVC in order to allow non-root users to access them. If you do not allow access by the sas user account, at a minimum, you will see permission errors. For more information about the sas user account, see User Accounts.

In your custom StorageClass definition, the UID and GID mount options must match the container process ID (which is 1001, sas, by default). Here is an example configuration:

```
cat << 'EOF' > StorageClass-RWX.yaml
kind: StorageClass
apiVersion: storage.k8s.io/v1
metadata:
  name: sas-azurefile
provisioner: kubernetes.io/azure-file
mountOptions:
  - dir_mode=0777
  - file_mode=0777
  - uid=1001
  - gid=1001
parameters:
  skuName: Standard_LRS
allowVolumeExpansion: true
EOF
kubectl apply -f StorageClass-RWX.yaml
```

Although Azure Kubernetes Service (AKS) provides a default azurefile StorageClass, the SAS testing organization found that this default StorageClass could not be modified to include the mount point options that SAS Viya requires. Instead, AKS quickly reverts the changes back to the default. Therefore, SAS recommends that you create a custom StorageClass in Azure environments. You can name it "sas-azurefile," as shown in this example. You can then configure this StorageClass when you update the base kustomization.yaml with the PVCs in your deployment that should use it. For more information, see Specify PersistentVolumeClaims to Use ReadWriteMany StorageClass.

If you want to use Azure File Services (azurefile) as a storage class for the PVC that is required for SAS Configurator for Open Source, an additional configuration step is required. When you provision the storage, verify that **`mfsymlinks`** is supplied as a mount option to the azurefile storage class. This option enables the update functionality.

File System and Shared Storage Requirements

The persistent volumes described previously are required for component and application data. However, various SAS Viya components also require a shared file system. A shared file system is required for multiple purposes, which include shared data storage and private user directories. A file server that uses the network file system (NFS) protocol is the minimum requirement.

If you are deploying a multi-tenant environment, each tenant will have its own CAS controller. A shared file system is required for all the path-based caslibs in each tenant. Mount the file system at `/opt/sas/tenant/config/data/cas/`. For *tenant*, substitute `viya` for a single tenant, or substitute the name of one of your tenants.

As you select a file system option, consider both disk size and the number of disks. Speed is directly related to the amount of disk space that is provisioned. It is also constrained by VM network and I/O limits.

In a Microsoft Azure environment, each VM instance type has a maximum input/output operations per second (IOPS) metric and a throughput metric, as well as a maximum number of disks that can be attached. The Microsoft Azure VM instances that SAS recommends are optimized for IOPS rather than for storage throughput. Striping can be used to increase storage IOPS and throughput, both of which are important for SAS Viya performance. The performance of attached disk configurations is constrained by VM limits.

The following options for shared file storage are recommended:

Table 7 *Storage Recommendations*

Cloud Provider	Standard Deployments	HA Deployments	Notes
Microsoft Azure	Azure Premium SSD Managed Disks with a shared file storage layer. When volumes are mounted to a VM instance that exports them as NFS, they have the equivalent of RWX accessMode. RAID 5 is recommended.	Azure NetApp Files	Both of these storage options are encrypted by default.
AWS	Elastic Block Store (EBS), mounted to a VM that exports with NFS. When volumes are mounted to a VM instance that exports them as NFS, they have the	Amazon Elastic File Share (EFS), with the following options: ■ Performance mode set to **Max I/O** ■ Throughput mode provisioned with	Installing a provisioner for EBS volumes, such as the nfs-provisioner or nfs-subdir-external-provisioner, is recommended.

Cloud Provider	Standard Deployments	HA Deployments	Notes
	equivalent of RWX accessMode.	**1024 MiB/s** or more	
GCP	Cloud Storage When volumes are mounted to a VM instance that exports them as NFS, they have the equivalent of RWX accessMode.	Cloud Filestore, SSD tier 5 x 375 GB local SSD (RAID0)	RAID 0 is recommended.

SAS also recommends using a consistent directory structure in your shared storage solution. Consistency is helpful for the following reasons:

- Multiple SAS Viya components require data storage.
- Some SAS offerings can also use a shared location for private user directories.
- Shared storage can serve as the location for persistent volumes to be provisioned in a consistent manner.
- SAS Viya uses some shared, open-source binary files.

To optimize your deployment, create the NFS directory structure that is described in the following table:

Table 8 *Shared File System Recommended Directory Structure*

Directory	Description
`/astores`	Location of the shared directory for ASTORES and ASTORE models.
`/bin`	Location for open-source companion software directories.
`/bin/nfsviyapython`	Location for your Python binary files.
`/bin/nfsviya-r`	Location for your R binary files.
`/data`	Location for SAS and CAS data.
`/homes`	Location for user private directories. As a best practice, create a subdirectory for each user of SAS Viya offerings.
`/pvs`	Location for persistent volumes that are provisioned using the file system.
`/permstore`	Location for the CAS server to store caslib authorization information.
`/backups`	Locations where SAS Viya backup files can be saved.
`/backups/cas`	

Directory	Description
`/backups/common`	
`/data-drivers/jdbc`	Location where JAR files for JDBC drivers are stored. If you add JDBC drivers for SAS Viya data access, save them in this location so that they are available to the deployment automatically.
`/quality-knowledge-base`	Location where SAS Quality Knowledge Base stores data.

File System Permissions

SAS Viya includes default fsgroup settings that enable file system access. When an fsgroup ID is set for a pod, any files that are written to a volume mounted by a container within that pod inherit that fsgroup as their group ID (GID). The fsgroup ID is the owner of the volume and of any files in that volume.

A Kubernetes administrator might want to change the default container security settings in a SAS Viya deployment by modifying settings in the podSpecs. For enhanced security, you might be able to remove the fsgroup settings. If your deployment is using a storage class provider that sets default file system permissions for persistent volumes to 777, the SAS Viya fsgroup settings are not required. You can use the remove-fsgroup-transformer resource to remove these settings during the deployment process. After you have downloaded deployment assets, locate the following README file for the instructions: `$deploy/sas-bases/examples/security/container-security/README.md` (for Markdown format) or `$deploy/sas-bases/docs/modify_container_security_settings.htm` (for HTML format).

However, the requirements for these settings are different for Red Hat OpenShift than they are for other deployment environments. If a pod is mounting an NFS volume, an SCC must be tied to the service account that enables it. For more information, see “SCCs and File System Permissions”.

Requirements for GPU Support

Deep learning capabilities are included with SAS Visual Machine Learning. The deep learning features are automatically installed. A graphics processing unit (GPU) is not required in your cluster in order to use deep learning features. However, a GPU provides additional functionality.

SAS Event Stream Processing also provides GPU support, but the requirements are different. To use a GPU with SAS Event Stream Processing, fulfill the requirements that are listed in “GPU Requirements for SAS Event Stream Processing” on page 22.

Supported GPU Configurations for Deep Learning

To enable deep learning with GPU functionality, here are additional requirements on Linux operating systems:

- [ ] A powerful, CUDA-capable GPU.

SAS has tested with GPUs that have the NVIDIA Compute Unified Device Architecture (CUDA).

Only the NVIDIA Quadro and Tesla product families are supported.

NVIDIA Pascal, Volta, Turing, and Ampere architectures are supported.

Note: If you have multiple GPUs installed on the same node, they must be of the same model.

- [ ] The NVIDIA display driver, version 440.33.01 or later. SAS recommends using the latest version.

  SAS also recommends enabling NVIDIA driver persistence mode at all times. For more information, see the Driver Persistence section of the NVIDIA deployment documentation.

  When you install the NVIDIA display driver on Linux, SAS recommends following the instructions in the NVIDIA CUDA Installation Guide for Linux.

  You can download the current drivers from http://www.nvidia.com/Download/index.aspx?lang=en-us.

- [ ] (For Microsoft Azure Only) N-Series VMs for the node pool that is labeled for CAS workloads. These VMs include GPU capabilities.

  The NVIDIA device plug-in must be installed and configured. For more information, see https://docs.microsoft.com/en-us/azure/aks/gpu-cluster. After you have downloaded and uncompressed the deployment assets, the README file in `$deploy/sas-bases/examples/gpu/` provides installation instructions.

- [ ] `/lib64` is the first path that is defined for the LD_LIBRARY_PATH environment variable on the server where the GPU is installed.

Run the following command on the machine where the GPU is installed in order to check the device type, the driver version, and the CUDA version:

```
nvidia-smi
```

Additional configuration is required in order to enable the SAS GPU Reservation Service on each CAS node. After you have downloaded and uncompressed the deployment assets, you can find a full set of instructions for enabling GPU support in `$deploy/sas-bases/examples/gpu/README.md` (for Markdown format) or `$deploy/sas-bases/docs/sas_gpu_reservation_service.htm` (for HTML).

GPU Requirements for SAS Event Stream Processing

SAS Event Stream Processing supports an optional GPU environment for high-powered analytics calculations, such as scoring with analytic store (ASTORE) files. A GPU enhances the deep learning functionality in SAS Event Stream Processing streaming analytics.

Here are the requirements for GPU support in SAS Event Stream Processing environments:

- [ ] GPU with NVIDIA Pascal or Volta architecture
- [ ] 10 GB or more of disk space

Resource Guidelines

The topics in this section provide recommendations for node scheduling, node sizing, and performance optimization.

Factors That Affect Resource Requirements

Several factors affect resource utilization by SAS Viya components, such as the following:

- The expected amount of data that SAS Viya users will process
- The expected number of concurrent users
- Whether an optional GPU is used in order to leverage the SAS Deep Learning feature
- Whether your CAS server implementation is SMP or MPP

  The CAS server can be deployed on a single node (SMP) or across several nodes (MPP). Distributing the CAS server across multiple nodes enables massively parallel processing (MPP). MPP CAS typically requires more resources.

Sizing for Migration

If you are migrating to this version of SAS Viya from a previous (3.*x*) version, the resource utilization of both deployments will be similar.

Here is an example: if your SAS Viya 3.5 deployment included three CAS server machines with 256 GB of RAM each, you should reserve nodes with a total of 768 GB of RAM in the Kubernetes cluster. If the Kubernetes administrator has allocated physical nodes of 64 GB of RAM each, you should label 12 of those nodes for CAS servers.

However, accounting for the typical 90% resource utilization, you can assume that the CAS controller and workers can reliably access only about 691 GB of RAM in this environment. Therefore, one additional node with 64 GB of RAM is recommended.

Sizing Recommendations for Microsoft Azure

A minimum of three node pools is required in your cluster: one default (system) node pool, where Kubernetes and other components that are not SAS Viya are deployed, and two user node pools for SAS Viya. Refer to "Default Node Pool Configuration" in *SAS Viya: Deployment Guide* for instructions on configuring the system node pool. One user node pool must be fully dedicated to the CAS server. The nodes in the CAS node pool require taints to prevent Kubernetes from scheduling non-CAS workloads on them. The other user node pool can host the remaining (non-CAS) component workloads. Consider creating five user node pools in order to accommodate the five workload classes.

The following table provides resource recommendations for representative SAS Viya product offerings. Derived from SAS performance testing, these estimates are for a *medium-sized* deployment, which was defined as 10 concurrent users that access SAS Viya user interfaces. In SAS testing simulations, these users manipulated data sets of 10–30 GB. VM instances all used an Intel Cascade Lake processor. The instance types are hyper-threaded, which means that two vCPUs are equivalent to one physical CPU core.

Table 9 *Resource Recommendations per Offering*

Offering	CAS Node Pool	System Node Pool	User Node Pools*
SAS Visual Analytics and SAS Data Preparation	RAM: 128 GB per instance CPU: 16 vCPUs per instance Example: Microsoft Azure E16ds_v4 Recommended minimum for CAS disk cache: 150 GB; ephemeral storage Example number of machines: 4	RAM: 64 GB CPU: 8 vCPUs Example: Microsoft Azure Standard_E8ds_v4 Example number of machines: 1	RAM: 64 GB CPU: 8 vCPUs per instance Example: Microsoft Azure E16ds_v4 Recommended minimum ephemeral storage for selected applications: 60 GB Example number of machines: 1 per node pool
SAS Visual Machine Learning	RAM: 128 GB per instance CPU: 16 vCPUs per instance Example: Microsoft Azure E16ds_v4 Recommended minimum for CAS disk cache: 600 GB; ephemeral storage Example number of machines: 8	RAM: 64 GB CPU: 8 vCPUs Example: Microsoft Azure Standard_E8ds_v4 Example number of machines: 1	RAM: 128 GB per instance CPU: 16 vCPUs per instance Example: Microsoft Azure E16ds_v4 Minimum disk: 2 x 128 GB Example number of machines: 1 per node pool
SAS Visual Data Science	RAM: 384 GB per instance CPU: 48 vCPUs per instance Example: Microsoft Azure E48ds_v4 Recommended minimum for CAS disk cache: 1200 GB Example number of machines: 9	RAM: 64 GB CPU: 8 vCPUs per instance Example: Microsoft Azure Standard_E8ds_v4 Example number of machines: 1	RAM: 128 GB per instance CPU: 48 vCPUs per instance Example: Microsoft Azure E16ds_v4 Minimum disk: 2 x 128 GB Example number of machines: 1 per node pool

* In addition to the CAS node pool, four user node pools are recommended in order to host the remaining SAS Viya workload classes.

Note: The additional resources that are recommended for SAS Visual Machine Learning are needed for compute workloads. The compute class includes pods that host SAS compute server instances, or that run batch jobs, or that do both.

The SAS workload node placement strategy suggests that you dedicate a node or a node pool on which Kubernetes schedules, or prefers scheduling, only SAS compute workloads. These workloads consist of pods that host SAS compute server instances, that run batch jobs, or that do both. This strategy enables more granular tuning when you want a specialized set of nodes to manage the compute workloads in the environment. SAS recommends selecting a machine with additional resources for the compute workload class. For more information about the workload node placement strategy, see "Plan the Workload Placement" in *SAS Viya: Deployment Guide*.

By default, updates are applied to your deployment using a strategy that avoids downtime. During each SAS Viya software update, the existing pods continue running until the new, updated pods have started up. Therefore, your cluster requires sufficient available resources to enable duplicates of each pod to run temporarily.

These guidelines do not attempt to account for all ordering scenarios, but instead are intended to illustrate typical software orders. SAS strongly recommends that you consult with a sizing expert to obtain an official hardware recommendation that is based on your requirements. To request sizing expertise, contact your SAS account representative. If you need assistance in determining your SAS account representative, send an email to contactcenter@sas.com.

Sizing Recommendations for AWS

A minimum of three AWS managed node groups is required in your cluster. One node group should be reserved for components that are not SAS Viya, and the remaining managed node groups are required for SAS Viya components. The reserved node group is referred to as the "default node pool" in this document. Refer to "Default Node Pool Configuration" in *SAS Viya: Deployment Guide* for instructions on configuring it. One managed node group must be fully dedicated to the CAS server. The nodes in the CAS node group require taints to prevent Kubernetes from scheduling non-CAS workloads on them. The other managed node group can be dedicated to the remaining (non-CAS) component workloads. Consider creating six managed node groups in order to accommodate the five workload classes and the default node pool.

The following table provides resource recommendations for a representative SAS Viya product offering, SAS Visual Machine Learning. Derived from SAS performance testing, these estimates are for a *medium-sized* deployment, which was defined as 10 concurrent users that access SAS Viya user interfaces. In SAS testing simulations, these users manipulated data sets of 10–30 GB. The EC2 VM instances used an Intel Cascade Lake series processor. In AWS, the number of vCPUs for an instance is indicated by a number in the name of the instance type, multiplied by 4. For example, the r5n.2xlarge instance type provides 8 vCPUs, the equivalent of 4 physical CPU cores.

Table 10 *Resource Recommendations for an Example Offering*

Workload Class	Resources
Default Node Pool	RAM: 16 GB CPU: 4 vCPUs Example: m5.xlarge Example number of machines: 1
CAS	RAM: 128 GB per instance CPU: 48 vCPUs per instance Example: r5dn.12xlarge Storage for CAS disk cache: 2 x 900 NVMe SSD

Workload Class	Resources
	Example number of machines: 3 (MPP) or 1 (SMP)
Connect	RAM: 128 GB per instance CPU: 16 vCPUs Example: m5n.4xlarge Example number of machines: 1
Compute	RAM: 384 GB per instance CPU: 48 vCPUs per instance Example: i3en.12xlarge Example number of machines: 2
Stateful or Stateless	RAM: 64 GB per instance CPU: 16 vCPUs per instance Example: m5n.4xlarge Example number of machines: 2 per node group

Note: CAS and compute nodes require high-performing storage.

By default, updates are applied to your deployment using a strategy that avoids downtime. During each SAS Viya software update, the existing pods continue running until the new, updated pods have started up. Therefore, your cluster requires sufficient available resources to enable duplicates of each pod to run temporarily.

These guidelines do not attempt to account for all ordering scenarios, but instead are intended to illustrate typical software orders. SAS strongly recommends that you consult with a sizing expert to obtain an official hardware recommendation that is based on your requirements. To request sizing expertise, contact your SAS account representative. If you need assistance in determining your SAS account representative, send an email to contactcenter@sas.com.

Sizing Recommendations for Google Cloud Platform

A minimum of three GCP node pools is required in your cluster. One node pool should be reserved for components that are not SAS Viya, and the remaining node pools are required for SAS Viya components. The reserved node pool is referred to as the “default node pool” in this document. Refer to “Default Node Pool Configuration” in *SAS Viya: Deployment Guide* for instructions on configuring it. One node pool must be fully dedicated to the CAS server. The nodes in the CAS node pool require taints to prevent Kubernetes from scheduling non-CAS workloads on them. The other node pool can be dedicated to the remaining (non-CAS) component workloads. Consider creating six node pools in order to accommodate the five workload classes and the default node pool.

The following table provides resource recommendations for a representative SAS Viya product offering, SAS Visual Machine Learning. Derived from SAS performance testing, these estimates are for a *medium-sized* deployment, which was defined as 7 concurrent users that access SAS Viya user interfaces. In SAS testing simulations, these users worked with table sizes of 76 GB and row counts of 10,000,000. The VM instances used an Intel Cascade Lake series processor.

Table 11 *Resource Recommendations for an Example Offering*

Workload Class	Resources
Default Node Pool	RAM: 32 GB CPU: 8 vCPUs Example: n2-standard-8 Example number of machines: 1
CAS	RAM: 384 GB CPU: 48 vCPUs Example: n2-highmem-48 Storage for CAS disk cache: 3 TiB; ephemeral Storage for CAS default data: 2.5 TiB, basic SDD (RAID 0) Example number of machines: 3 (MPP) or 1 (SMP)
Connect	RAM: 128 GB per instance CPU: 16 vCPUs Example: n2-highmem-16 Example number of machines: 1
Compute	RAM: 384 GB per instance CPU: 48 vCPUs per instance Example: n2-highmem-48 Example number of machines: 1
Stateful or Stateless	RAM: 64 GB per instance CPU: 16 VCPUs per instance Example: n2-standard-16 Recommended minimum ephemeral storage for selected applications: 60 GB Example number of machines: 2

Note: CAS and compute nodes require high-performing storage.

By default, updates are applied to your deployment using a strategy that avoids downtime. During each SAS Viya software update, the existing pods continue running until the new, updated pods have started up. Therefore, your cluster requires sufficient available resources to enable duplicates of each pod to run temporarily.

These guidelines do not attempt to account for all ordering scenarios, but instead are intended to illustrate typical software orders. SAS strongly recommends that you consult with a sizing expert to obtain an official hardware recommendation that is based on your requirements. To request sizing

expertise, contact your SAS account representative. If you need assistance in determining your SAS account representative, send an email to contactcenter@sas.com.

Sizing Recommendations for Red Hat OpenShift

Note: It is helpful to understand SAS Viya workloads so that you can manage them effectively. The workload placement strategy that is described in *SAS Viya: Deployment* might not be appropriate for a SAS Viya deployment in a shared OpenShift cluster on VMWare, where high node utilization is the goal. For each installation, you should assess the level of isolation that is required for SAS Viya workloads from other applications in the cluster. However, the label "workload.sas.com/class=compute" that is described in that section of the documentation is required if your order includes SAS Workload Management.

Red Hat recommends allocating three nodes for the Kubernetes control plane. These nodes can have 4 CPU cores with 16 GB of RAM. They support load balancing, service discovery, batch execution, and other tasks. Nodes that are not part of the control plane and that do not have taints or labels for a SAS workload class are referred to as “default nodes” in this guide. Review the Red Hat documentation to understand the control plane machine config pool.

If your SAS Viya installation uses the CAS server, SAS also recommends that you dedicate at least one node to CAS by labeling and tainting a node in the worker machine config pool. If you are deploying MPP CAS, label and taint additional nodes. Another option is to create a custom pool for CAS nodes in order to have dedicated auto scaling. In addition, SAS strongly recommends labeling one node in the worker machine config pool for compute workloads. The remaining nodes in the worker machine config pool will be targets for the rest of the SAS Viya workloads.

For best performance, select machines with the highest available I/O throughput levels. For all types of required storage, SAS recommends sequential I/O bandwidth of 90–120 MB per second, per CPU core. Set up storage with an understanding of the I/O bandwidth performance that your machines can achieve. If applicable for your machines, set server power settings to maximum.

The following table provides resource recommendations for a representative SAS Viya product offering, SAS Visual Machine Learning. Derived from SAS performance testing, these estimates are for a *medium-sized* deployment, which was defined as 7 concurrent users that access SAS Viya user interfaces. In SAS testing simulations, these users worked with table sizes of up to 50 GB. All machines used an Intel Xeon Gold processor. CPUs all used hyperthreading.

Table 12 *Resource Recommendations for the Worker Node Pool in an Example Offering*

Workload Class	Resources per Node
Default	RAM: 32 GB CPU: 4 vCPUs Example number of nodes: 2–4 Additional nodes in the worker machine config pool can be used to host SAS Viya pods.
CAS	RAM: 192 GB (assuming 3 nodes) CPU: 24 vCPUs Storage for CAS disk cache: 300 GB; use high-performing storage

Workload Class	Resources per Node
	Storage for CAS default data: 2 x 400, SSD (data redundancy is recommended) Example number of nodes: 3 (MPP) or 1 (SMP)
Connect	RAM: 16 GB CPU: 2 vCPUs Example number of nodes: 1
Compute	RAM: 32 GB CPU: 8 vCPUs Example number of nodes: 1
Stateful	RAM: 32 GB or more CPU: 8 VCPUs Recommended minimum ephemeral storage: 60 GB Example number of nodes: 1
Stateless	RAM: 80 GB or more CPU: 8 VCPUs Example number of nodes: 1

Note: CAS and compute nodes require high-performing storage.

By default, updates are applied to your deployment using a strategy that avoids downtime. During each SAS Viya software update, the existing pods continue running until the new, updated pods have started up. Therefore, your cluster requires sufficient available resources to enable duplicates of each pod to run temporarily.

These guidelines do not attempt to account for all ordering scenarios, but instead are intended to illustrate typical software orders. SAS strongly recommends that you consult with a sizing expert to obtain an official hardware recommendation that is based on your requirements. To request sizing expertise, contact your SAS account representative. If you need assistance in determining your SAS account representative, send an email to contactcenter@sas.com.

CAS Server Resources

A CAS server consists of either a single node (SMP), or a set of nodes that include one controller, optionally one backup controller, and multiple workers (MPP). (Although a one-worker MPP configuration is supported, it is not an efficient allocation of resources.) The nodes that you select for the CAS server have the following requirements:

- Guaranteed quality of service (QoS).

CAS pods must execute with guaranteed QoS in order to avoid shutdowns when memory resources are limited. Kubernetes QoS requires that the requests and limits be set to equal numbers for each container in the pod.

- Adequate RAM and CPU resources.

  By default, auto-resourcing is applied: the CAS operator determines the amount of RAM for your deployment based on available RAM on the nodes where CAS workloads are running.

  The file `sas-bases/examples/cas/configure/cas-manage-cpu-and-memory.yaml` can be applied to the kustomization.yaml file if you instead want to manually specify resource requests and limits.

- Storage for the CAS disk cache.

  Every CAS pod requires a mount for this caching space. The mount is an empty directory that points to `/cas/cache`. It is configured automatically by the deployment process.

  Performance improves if you use local storage, such as an SSD volume that is memory-mapped to provide overflow capacity for data that the CAS server processes in memory. Configure the CASENV_CAS_DISK_CACHE environment variable to point to it. For more information, see Tune CAS_DISK_CACHE.

  In a cloud environment, ephemeral storage is recommended. The required amount depends on the size of table data and how the CAS server accesses it.

- Storage for required CAS PVCs.

  These PVCs are described in “Persistent Storage Volumes, PersistentVolumeClaims, and Storage Classes” on page 15.

- Fully dedicated nodes that do not share resources with any other SAS Viya components.

  To create these dedicated resources, apply taints and labels to the appropriate nodes in the CAS node pool. This recommendation ensures that CAS QoS can be achieved. The recommended node taints prevent Kubernetes from scheduling workloads that could compete for resources with CAS. See Plan Workload Placement for more information.

  If resources on CAS-dedicated nodes are exhausted, CAS might be scheduled to nodes outside of the CAS node pool. If external scheduling occurs, contact your SAS account representative to address issues with CAS node pool sizing. To avoid this situation, SAS recommends that you add taints to other nodes in the cluster in conformance with the SAS workload placement strategy.

  IMPORTANT Tainting all nodes in the cluster limits workloads from components that are not SAS components from being scheduled to tainted nodes. Otherwise, adjustments to manifests are necessary for workloads that are not SAS components.

- (Optional) A secondary (backup) CAS controller to enable failover.

  In a multi-tenant deployment, you can have multiple backup controllers (one per tenant). MPP CAS is required. For more information, see “Add a Backup Controller for MPP CAS” in *SAS Viya: Deployment Guide*.

- In a multi-tenant deployment, additional nodes with guaranteed QoS that are labeled for CAS

  Each tenant requires its own dedicated CAS server. Each CAS pod must be placed on its own node with the required nodeAffinity settings that are described in Plan Workload Placement.

SAS Programming Runtime Environment Requirements

The SAS Programming Run-Time Environment, which includes SAS Compute Server, SAS/CONNECT server, and SAS Batch Server, produces workloads in the compute workload class. These components must be able to create temporary files and data sets in a storage volume.

By default, sas-programming-environment pods are backed by an emptyDir volume named **viya**. This volume is mounted automatically and uses storage on the node. Using the default emptyDir volume is not recommended because SAS programming components can consume large amounts of storage quickly and cause nodes to shut down.

One example of a critical directory that is allocated on the **viya** volume to support SAS Programming Run-Time Environment components is SASWORK, which can rapidly outgrow a local storage volume. High-performing storage is important in order to avoid degraded SAS Viya performance. Ephemeral storage is an option for SASWORK, but you must prevent SAS processes from consuming disk space that is required by the node's kubelet. If your SAS Batch Server will be running jobs in SAS checkpoint/restart mode, persistent external storage is required for the corresponding **viya** volume. This requirement enables the checkpoint information that is stored in SASWORK to be saved in the case of a node failure or job preemption.

For information about modifying storage classes for SAS Programming Run-Time servers, see "Configure External Storage Class for SAS Programming Run-Time Environment" in *SAS Viya: Deployment Guide*.

Note: The name **viya** is required for the volume that supports SAS Programming Run-Time Environment components.

Virtual Machine Recommendations for Microsoft Azure

The following table provides guidelines to help you set up an appropriate cloud environment for your Kubernetes deployment of SAS Viya. Keep in mind that these are minimum recommended allocations. Monitoring workload levels and available storage is recommended in order to maintain adequate resources for the cluster.

Table 13 *Cloud Components and Minimum Sizing*

Component	Description	Size
Nodes	Select virtual machines with good I/O, such as the D(s/ds)v3 Series machines. These instances provide hyper-threading and up to 3.5 GHz. Use the following general guidance: ■ Select the E16ds_v4 or equivalent for the CAS server. ■ Select Standard_D16s_v3 for most SAS Viya orders.	One large deployment: 5 or more nodes Two or more large deployments: 7 or more nodes With auto-scaling, the following are recommended: ■ 2 or more nodes in each of the SAS Viya workload node pools ■ 1 node in the system node pool

Component	Description	Size
	■ Select Standard_D8s_v3 for SAS Viya orders with fewer offerings (1–2). ■ Select the auto-scaling option for node pools where possible.	
Disk Size	In Microsoft Azure, each node includes ephemeral storage.	200 GB is the recommended minimum. The default 100 GB disk size in Azure was not adequate in SAS testing.
Database Connections	A PostgreSQL database is required to host SAS Infrastructure Data Server. You can use the default *internal* database that is deployed automatically, or you can supply your own *external* database. For more information, see "Internal versus External PostgreSQL Instances".	A typical deployment with two or three SAS Viya offerings requires a minimum of 1024 database user connections to avoid running out of connections. The internal PostgreSQL instance is set to 1280 connections by default. An external PostgreSQL server should support maximum connections and maximum prepared transactions of at least 1024. The requirements for a multi-tenant deployment might differ. For more information, see "PostgreSQL Requirements for a Multi-Tenant Deployment" on page 66.
Database Storage	A single PostgreSQL server can support multiple SAS Viya deployments. You define a separate database for each deployment.	A typical deployment with two or three SAS Viya offerings requires approximately 1–2 GB for the deployment process. If a single PostgreSQL server supports multiple SAS Viya deployments that are typical in size, use the following guidance to determine storage space: ■ For 1–2 deployments, use at least 128 GB of storage space. ■ For 3–4 deployments, use at least 512 GB of storage space.

Consider the following guidelines as you select an image type for the VM instances in your cluster:

- Select VM instance types with Intel chips. If the name of the Azure image contains the letter “a,” the VM is based on an AMD chip (such as the Dav4 and Dasv4-series). Multiple SAS Viya processes performed significantly faster on Intel chips in SAS testing.
- For the v4 type of image, if the name contains the letter “d” indicating “disk,” the node includes a local temp disk. The CAS server requires ephemeral storage.
- If the node is intended to host a pod that requires a PVC, SAS recommends that you select a VM image that supports premium storage. Premium storage is indicated by the letter “s” in the image name. For more information, see “Storage Requirements”.
- SAS Viya performs better on nodes with a high memory-to-CPU ratio. The Microsoft Azure E series images have a higher memory-to-CPU ratio than the D series.

Virtual Machine Recommendations for AWS

The following table provides guidelines to help you set up an appropriate cloud environment for your EKS deployment of SAS Viya. Keep in mind that these are minimum recommended allocations. Monitoring workload levels and available storage is recommended in order to maintain adequate resources for the cluster.

Table 14 *Cloud Components and Minimum Sizing*

Component	Description	Size
Nodes	Select virtual machines with good I/O, such as the m5n, r5d, or i3 series machines. These instances provide hyper-threading and 3.1–3.5 GHz processors. The i3 series processors are slower but have more ephemeral storage and are suitable for compute workloads. Use the following general guidance: ■ The r5(d)n, m5(d)n, i3en families have higher network throughput than the r5(d), m5(d) and i3 series. ■ Select the equivalent of an r5dn.large for the CAS server. ■ Select the equivalent of an i3en.large for the compute server. ■ For a medium-sized SAS Viya order with multiple offerings, select m5n.large or the equivalent for the remaining nodes.	One large deployment: 5 or more nodes Two or more large deployments: 7 or more nodes With auto-scaling, the following are recommended: ■ 2 or more nodes in each of the SAS Viya workload node groups ■ 1 node in the default node group (default node pool)

Component	Description	Size
	■ Select the auto-scaling option for node groups where possible.	
Disk Size	In AWS, each node includes ephemeral storage. A “d” in the name of the VM instance type indicates that it includes NVMe SSD storage, while an “e” indicates additional storage.	200 GB is the recommended minimum. The default instance store size in AWS is typically adequate.
Database Connections	A PostgreSQL database is required to host SAS Infrastructure Data Server. You can use the default *internal* database that is deployed automatically, or you can supply your own *external* database. For more information, see “Internal versus External PostgreSQL Instances”.	A typical deployment with two or three SAS Viya offerings requires a minimum of 1024 database user connections to avoid running out of connections. An external PostgreSQL server should support maximum connections and maximum prepared transactions of at least 1024. The requirements for a multi-tenant deployment might differ. For more information, see “PostgreSQL Requirements for a Multi-Tenant Deployment” on page 66.
Database Storage	A single PostgreSQL server can support multiple SAS Viya deployments. You define a separate database for each deployment.	A typical deployment with two or three SAS Viya products requires approximately 1–2 GB for the deployment process. If a single PostgreSQL server supports multiple SAS Viya deployments that are typical in size, use the following guidance to determine storage space: ■ For 1–2 deployments, use at least 128 GB of storage space. ■ For 3–4 deployments, use at least 512 GB of storage space.

Consider the following guidelines as you select instance types for the VM instances in your cluster:

- Select VM instance types with Intel chips. Multiple SAS Viya processes performed significantly faster on Intel chips in SAS testing.
- SAS Viya performs better on nodes with a high memory-to-CPU ratio.

Virtual Machine Recommendations for Google Cloud Platform

The following table provides guidelines to help you set up an appropriate cloud environment for your Kubernetes deployment of SAS Viya. Keep in mind that these are minimum recommended allocations. Monitoring workload levels and available storage is recommended in order to maintain adequate resources for the cluster.

Table 15 *Cloud Components and Minimum Sizing*

Component	Description	Size
Nodes	Select virtual machines with additional memory, such as the N2 high-memory machine type. These instances provide 6.5 GB of memory per vCPU. Use the following general guidance: ■ Select the n2-highmem-48 or equivalent for the CAS server. ■ Select the n2-standard-16 for stateful or stateless workloads. ■ Select the auto-scaling option for node pools where possible.	One large deployment: 5 or more nodes Two or more large deployments: 7 or more nodes With auto-scaling, the following are recommended: ■ 2 or more nodes in each of the workload node pools ■ 1 node in the default node pool
Disk Size	Basic SSD was found to be adequate in SAS testing. Use Google File Store with the “Basic SSD” service tier for persistent volume storage. For the CAS disk cache, use 5 or more local SSD (RAID0) persistent disks with 375 GB of space.	200 GB is the recommended minimum.
Database Connections	A PostgreSQL database is required to host SAS Infrastructure Data Server. You can use the default *internal* database that is deployed automatically, or you can supply your own *external* database. For more information, see “Internal versus External PostgreSQL Instances”.	A typical deployment with two or three SAS Viya product offerings requires a minimum of 1024 database user connections to avoid running out of connections. The internal PostgreSQL instance is set to 1280 connections by default. An external PostgreSQL server should support maximum connections and maximum prepared transactions of at least 1024.

Component	Description	Size
		The requirements for a multi-tenant deployment might differ. For more information, see "PostgreSQL Requirements for a Multi-Tenant Deployment" on page 66.
Database Storage	A single PostgreSQL server can support multiple SAS Viya deployments. You define a separate database for each deployment.	A typical deployment with two or three SAS Viya offerings requires approximately 1–2 GB for the deployment process. If a single PostgreSQL server supports multiple SAS Viya deployments that are typical in size, use the following guidance to determine storage space: ■ For 1–2 deployments, use at least 128 GB of storage space. ■ For 3–4 deployments, use at least 512 GB of storage space.

Consider the following guidelines as you select an image type for the VM instances in your cluster:

- Select VM instance types with Intel chips. Multiple SAS Viya processes performed significantly faster on Intel chips in SAS testing.
- GCP zonal or regional persistent disks can yield improved IOPS and throughput on instances with more vCPUs, such as the n2-highmem-48.
- SAS Viya performs better on nodes with a high memory-to-CPU ratio, such as the GCP N2 machine type. The extended memory feature provides more memory per CPU.

Virtual Machine Recommendations for Red Hat OpenShift

The following table provides guidelines to help you set up an appropriate on-premises environment for your vCenter deployment of SAS Viya. Keep in mind that these are minimum recommended allocations. Monitoring workload levels and available storage is recommended in order to maintain adequate resources for the cluster.

Table 16 *Components and Minimum Sizing*

Component	Description	Size
Nodes	Select machines with good I/O and hyper-threading. Select a more powerful machine for the CAS server.	One large deployment: 5 or more nodes Two or more large deployments: 7 or more nodes The following are recommended:

Component	Description	Size
		■ 2 or more nodes in the worker machine config pool for each SAS Viya workload class ■ 3 nodes in the control-plane machine config pool (a Red Hat requirement)
Disk Size	The latest server generations are preferable because they are optimized for applications that, like SAS Viya, require faster CPUs, better local disk performance, and more RAM.	200 GB is the recommended minimum.
Database Connections	A PostgreSQL database is required to host SAS Infrastructure Data Server. You can use the default *internal* database that is deployed automatically, or you can supply your own *external* database. For more information, see "Internal versus External PostgreSQL Instances".	A typical deployment with two or three SAS Viya product offerings requires a minimum of 1024 database user connections to avoid running out of connections. The internal PostgreSQL instance is set to 1280 connections by default. An external PostgreSQL server should support maximum connections and maximum prepared transactions of at least 1024.
Database Storage	A single PostgreSQL server can support multiple SAS Viya deployments. You define a separate database for each deployment.	A typical deployment with two or three SAS Viya offerings requires approximately 1–2 GB for the deployment process. If a single PostgreSQL server supports multiple SAS Viya deployments that are typical in size, use the following guidance to determine storage space: ■ For 1–2 deployments, use at least 128 GB of storage space. ■ For 3–4 deployments, use at least 512 GB of storage space.

Consider the following guidelines as you select machines for your cluster:

- Use nodes with Intel chips. Multiple SAS Viya processes performed significantly faster on Intel chips in SAS testing.
- For nodes that require local storage PVCs, in order to meet I/O bandwidth requirements, SAS recommends that you select nodes with high-performance storage options, such as SSD/NVMe, SAN-attached storage, or similar options. For more information, see "Storage Requirements".

- SAS Viya generally performs better on machines with a high memory-to-CPU ratio.

Data Source Requirements

Data Source Requirements

Your software order included SAS/ACCESS products to support your data sources. Data access requires additional customization of your Kubernetes deployment and in some cases also requires installation of other software.

Individual SAS Viya product offerings might support a subset of these data sources. Refer to the lists of data sources that are supported by individual offerings. Then consult the list of requirements for the SAS/ACCESS product that supports your data source for additional system requirements that apply to your environment.

Supported Data Sources

The following external data sources are supported by most SAS Viya offerings, including SAS Visual Analytics, SAS Visual Statistics, SAS Visual Machine Learning, SAS Visual Data Science, SAS Visual Data Science Decisioning, SAS Visual Forecasting, and SAS Visual Text Analytics:

- Amazon Redshift
- Google BigQuery
- Greenplum
- Hadoop with Hive
- IBM DB2
- IBM Netezza
- Impala
- Data sources accessible with a Java Database Connectivity (JDBC) driver
- Microsoft SQL Server
- MongoDB
- MySQL
- Data sources accessible with an ODBC driver
- Oracle
- PC files
- PI System
- PostgreSQL
- Salesforce
- SAP ASE

- SAP HANA
- SAP R/3
- Snowflake
- Spark
- Teradata
- Vertica
- Yellowbrick

Data Sources for SAS Micro Analytic Service

SAS Micro Analytic Service supports only the following subset of the data sources that SAS Viya supports:

- Microsoft Azure PostgreSQL
- Microsoft Azure SQL Database via ODBC
- Microsoft SQL Server via ODBC
- Oracle
- PostgreSQL

General Requirements for SAS/ACCESS

Before you start the deployment, collect the third-party libraries and configuration files that are required for your data sources. Examples of these requirements include the following:

- Third-party drivers
- ODBC drivers
- JDBC drivers
- Hadoop configuration files

When you have collected these files, place them on storage that is accessible to your Kubernetes cluster. This storage could be a mount or a storage device with a persistent volume configured.

SAS recommends organizing your software in a consistent manner on your mount or storage device. Take note of the details for your specific storage solution, as well as the paths to the configuration files within it. You will need this information before you start the deployment.

Requirements for SAS/ACCESS Interface to Amazon Redshift

SAS/ACCESS Interface to Amazon Redshift includes SAS Data Connector to Amazon Redshift.

The required client software is included with your SAS Viya installation. Using a Data Source Name (DSN) to connect to Amazon Redshift requires post-installation configuration of your Kubernetes deployment.

Requirements for SAS/ACCESS Interface to DB2

SAS/ACCESS Interface to DB2 includes SAS Data Connector to DB2.

IBM DB2 Connect™ must also be licensed if you plan to connect to IBM DB2 databases that are running on AS/400, VSE, VM, MVS, and z/OS systems. The following DBMS products are supported:

- IBM DB2 version 11.5 or later
- IBM Integrated Analytics System (IIAS)
- Client utilities for IBM DB2 version 11.5 or later

SAS recommends installing the latest FixPack on the client and server.

Requirements for SAS/ACCESS Interface to Google BigQuery

SAS/ACCESS Interface to Google BigQuery includes SAS Data Connector to Google BigQuery.

Required client software is included with your SAS Viya installation. No additional software is required.

Requirements for SAS/ACCESS Interface to Greenplum

SAS/ACCESS Interface to Greenplum includes SAS Data Connector to Greenplum.

SAS/ACCESS Interface to Greenplum supports Greenplum Database versions 6.5 or later, and the required client software is included with your SAS Viya installation. Using a Data Source Name (DSN) to connect to Greenplum requires post-installation configuration of your Kubernetes deployment.

Requirements for SAS/ACCESS Interface to Hadoop

SAS/ACCESS Interface to Hadoop includes SAS Data Connector for Hive.

Hive 1.1 or later is required.

SAS/ACCESS Interface to Hadoop supports the following Hadoop distributions:

- Cloudera CDH 6.2

- Cloudera Data Platform (CDP) 7.1 (Public and Private Cloud)
- Hortonworks HDP 3.1
- Amazon Web Services EMR 5.13
- Microsoft Azure HDInsight 3.6 and 4

The SAS policy that applies to alternative releases or distributions of Hadoop is documented on the SAS Support website.

You must run the Hadoop Tracer Script for SAS Viya in order to install required JAR files that enable the CAS server to access files in Hadoop. The script has the following requirements:

- The user who is running the script must have authorization to issue HDFS and Hive commands.
- If your Hadoop deployment is secured with Kerberos, obtain a Kerberos ticket for the user before running the script.
- Python and the strace Linux library must be installed on the Hadoop cluster. Install them from the package repositories for your Linux distribution if necessary.

  Python 2.6 or later is required.

Requirements for SAS/ACCESS Interface to Impala

SAS/ACCESS Interface to Impala includes SAS Data Connector to Impala.

SAS/ACCESS Interface to Impala supports Impala Server 3.2 or a later version. It also supports the ODBC Driver for Impala, version 2.6.13 or a later version.

SAS/ACCESS Interface to Impala is supported on Cloudera Data Platform (CDP) Public and Private Cloud.

In order to take advantage of SAS/ACCESS bulk loading functionality, you must run the Hadoop Tracer Script. The script installs required JAR files that enable the SAS Viya servers to access files in Hadoop.

Requirements for SAS In-Database Technologies

SAS In-Database Technologies is a technology bundle that is included with multiple SAS Viya product offerings. It includes products that support distributed data sources and have distinct system requirements. To use SAS In-Database Technologies, SAS Embedded Process must be installed.

Requirements for SAS In-Database Technologies for Hadoop

SAS In-Database Technologies for Hadoop supports the following distributions of Hadoop:

- Amazon Web Services EMR 5.30
- Hortonworks HDP 3.1
- Cloudera CDH 6.2
- Cloudera CDP 7.1 Private Cloud
- Cloudera CDP 7.2 Public Cloud

Execution of SAS Embedded Process for Hadoop on the Spark platform requires Spark 2. Spark 2 is supported on HDP 3.1, CDH 6.2, CDP 7.1, and CDP 7.2. Spark is not supported on EMR 5.30 or MapR 6.1.

Requirements for SAS In-Database Technologies for Spark

SAS In-Database Technologies for Spark supports the following data sources with Spark distributions:

- Databricks 6.*x* (with Spark 2.4) for Microsoft Azure or Amazon Web Services
- Databricks 7.*x* and 9.*x* (both with Spark 3.*x*) for Microsoft Azure or Amazon Web Services
- Microsoft Azure Synapse Analytics (with Spark 2.4 and 3.1)

Requirements for SAS In-Database Technologies for Teradata

SAS In-Database Technologies for Teradata supports the following products:

- Teradata Vantage Advanced SQL Engine version 16.20 or later
- Teradata CLIv2 client libraries, TTU 16.20 or later for Linux (64-bit libraries)

SAS In-Database Technologies for Teradata also supports Teradata Vantage on the following cloud platforms:

- Amazon Web Services
- Google Cloud Platform
- Microsoft Azure
- Teradata Cloud
- VMware

Requirements for SAS/ACCESS Interface to JDBC

SAS/ACCESS Interface to JDBC includes SAS Data Connector to JDBC. SAS/ACCESS Interface to JDBC enables access to relational databases by means of SQL and the Java Database Connectivity (JDBC) API.

A JDBC driver is required for the data source from which you want to access data. JDBC drivers are available from DBMS vendors and other third-party JDBC driver developers.

Requirements for SAS/ACCESS Interface to Microsoft SQL Server

SAS/ACCESS Interface to Microsoft SQL Server includes SAS Data Connector to Microsoft SQL Server.

SAS/ACCESS Interface to Microsoft SQL Server supports the following products:

- Amazon RDS Microsoft SQL Server (Microsoft SQL Server 2017 or later)
- Google Cloud Platform Cloud SQL for SQL Server version 2017 or later
- Microsoft SQL Server 2017 or later
- Microsoft Azure SQL Database
- Microsoft Azure SQL Database Managed Instance
- Microsoft Azure SQL Server Big Data Clusters
- Microsoft Azure Synapse

Required client software is included with your SAS Viya installation. Using a Data Source Name (DSN) to connect requires some post-installation configuration of your Kubernetes deployment.

A DataDirect ODBC driver that is included with SAS/ACCESS to Microsoft SQL Server enables you to use a Microsoft SQL Server data source with SAS Micro Analytic Service.

Requirements for SAS/ACCESS Interface to MongoDB

SAS/ACCESS Interface to MongoDB includes SAS Data Connector to MongoDB.

SAS/ACCESS Interface to MongoDB requires the MongoDB C Driver version 1.17 or later ("libmongoc," the official client library for C applications). You can obtain the latest MongoDB C driver from the following website: http://mongoc.org/.

SAS/ACCESS Interface to MongoDB supports the following databases:

- MongoDB version 4.0 or later
- MongoDB Atlas tiers M10 and higher, for SAS Viya Stable 2021.1.6 or Long-Term Support 2021.2 and later.

Requirements for SAS/ACCESS Interface to MySQL

SAS/ACCESS Interface to MySQL includes SAS Data Connector to MySQL.

SAS/ACCESS Interface to MySQL requires MySQL Client version 5.6 or later.

The following DBMS products are supported:

- Amazon Aurora (MySQL engine version 5.6 or later)
- Amazon RDS MariaDB (engine version 10.1 or later)
- Amazon RDS MySQL (engine version 5.6 or later)
- Azure Database for MySQL (engine version 5.6 or later)
- Google Cloud SQL for MySQL (engine version 5.6 or later)
- MySQL Server version 5.6 or later
- MariaDB 10.1 or later

- SingleStore (MemSQL) 6.0 or later
- Oracle MySQL Database

Requirements for SAS/ACCESS Interface to Netezza

SAS/ACCESS Interface to Netezza includes SAS Data Connector to Netezza.

SAS/ACCESS Interface to Netezza requires the IBM Netezza ODBC driver from IBM. To obtain the appropriate IBM Netezza ODBC driver, contact IBM Technical Support at (877) 426-6006 or visit the IBM Fix Central website: http://www.ibm.com/support/fixcentral.

SAS Viya supports the following DBMS products:

- IBM Netezza 7.2.1 or later
- IBM Netezza Performance Server 11.2.0.0 and 11.2.1.*x*

Requirements for SAS/ACCESS Interface to ODBC

SAS/ACCESS Interface to ODBC includes SAS Data Connector to ODBC. SAS/ACCESS Interface to ODBC enables access to multiple data source types by means of a generic ODBC driver.

Before you can use SAS Viya with ODBC, an ODBC driver is required for the data source from which you want to access data. ODBC drivers are often available from DBMS vendors and other third-party ODBC driver developers. Your ODBC driver must comply with the ODBC 3.5 (or later) specification.

Note: The ODBC driver that you select might require additional DBMS software in order to enable network access.

Requirements for SAS/ACCESS Interface to Oracle

SAS/ACCESS Interface to Oracle (on SAS Viya) includes SAS Data Connector to Oracle.

SAS/ACCESS Interface to Oracle requires the Oracle client 12c or later (64-bit libraries).

SAS Viya supports the following Oracle instances:

- Amazon RDS Oracle 12c or later
- Oracle Cloud Platform 12c or later
- Oracle Database 12c or later

Obtain the path to the volume on the Oracle server to which you want to point your SAS Viya deployment. This information is used later in the deployment.

Requirements for SAS/ACCESS Interface to PC Files

SAS/ACCESS Interface to PC Files includes SAS Data Connector to PC Files.

SAS/ACCESS Interface to PC Files enables access to the following file formats:

- `.jmp`
- `.spss`
- `.stata`
- `.xlsx` or `.xls`

Additional Microsoft file formats can be accessed via the PCFILES LIBNAME engine, which is included with SAS/ACCESS Interface to PC Files.

No additional software is required.

Requirements for SAS/ACCESS Interface to the PI System

SAS/ACCESS Interface to the PI System uses the PI System Web API, which is HTTPS-based and RESTful. No PI System client software is required to be installed on the nodes where SAS is running. However, the PI System Web API (PI Web API 2019 or later) must be installed and activated on the host machine from which the user connects.

Requirements for SAS/ACCESS Interface to PostgreSQL

SAS/ACCESS Interface to PostgreSQL (on SAS Viya) includes SAS Data Connector to PostgreSQL.

SAS/ACCESS Interface to PostgreSQL supports:

- Amazon Aurora (PostgreSQL engine version 9.6 or later)
- Amazon RDS PostgreSQL (engine version 9.6 or later)
- Azure Database for PostgreSQL (engine version 9.6 or later)
- EnterpriseDB PostgreSQL version 9.6 or later
- Google Cloud Platform Cloud SQL for PostgreSQL (engine version 9.6 or later)
- PostgreSQL Database version 9.6 or later

Required client software is included with your SAS Viya installation. Using a Data Source Name (DSN) to connect requires some post-installation configuration of your Kubernetes deployment.

Requirements for SAS/ACCESS Interface to Salesforce

SAS/ACCESS Interface to Salesforce includes SAS Data Connector to Salesforce.

SAS/ACCESS Interface to Salesforce requires a Salesforce user account that has API access enabled. SAS/ACCESS Interface to Salesforce supports Salesforce API access, version 46.0 or later.

Requirements for SAS/ACCESS Interface to SAP ASE

SAS/ACCESS Interface to SAP ASE requires SAP ASE (formerly Sybase) Open Client SDK, Release 15.7 or later (64-bit libraries).

The database administrator must install two SAP ASE (Sybase) stored procedures on the target SAP server. These files are available in a compressed TGZ archive for download from the SAS Support site at https://support.sas.com/downloads/package.htm?pid=2458.

Requirements for SAS/ACCESS Interface to SAP HANA

SAS/ACCESS Interface to SAP HANA includes SAS Data Connector to SAP HANA.

SAS/ACCESS Interface to SAP HANA supports SAP HANA SPS 11 Server or a later version and requires SAP HANA ODBC Client (64-big) for SPS 11 or later.

Requirements for SAS/ACCESS Interface to R/3

SAS/ACCESS Interface to SAP R/3 supports SAP NetWeaver 7.0 (Application Server ABAP) or later and requires the 64-bit SAP NetWeaver RFC Library, Release 7.20 or later, which is provided by SAP AG.

Requirements for SAS/ACCESS Interface to Snowflake

SAS/ACCESS Interface to Snowflake includes SAS Data Connector to Snowflake. The required client software is also included with your SAS Viya installation.

Using a Data Source Name (DSN) to connect to Snowflake requires post-installation configuration of your Kubernetes deployment.

Requirements for SAS/ACCESS Interface to Spark

SAS/ACCESS Interface to Spark includes SAS Data Connector to Spark and supports Spark server 3.1 or later.

SAS/ACCESS to Spark requires the Spark Thrift Server. A vendor-supplied JAR file is required for Databricks access.

For other Spark configurations, you must run the Hadoop Tracer Script in order to install required JAR files that enable the SAS Viya servers to access files in Hadoop.

SAS/ACCESS Interface to Spark supports the following data sources:

- Hortonworks HDP 3.1 with the required Spark client JAR files
- Azure Databricks 8.3 or later with the latest Databricks JDBC driver
- Databricks on Google Cloud 8.3 or later with the latest Databricks JDBC driver
- Databricks on Amazon Web Services (AWS) 8.3 or later with the latest Databricks JDBC driver

Requirements for SAS/ACCESS Interface to Teradata

SAS/ACCESS Interface to Teradata includes SAS Data Connector to Teradata.

SAS/ACCESS Interface to Teradata requires Teradata CLIv2 client libraries, TTU 16.20 or later, and supports Teradata Vantage SQL Engine version 16.20 in addition to the following platforms:

- Teradata IntelliCloud
- Teradata Vantage on Amazon Web Services
- Teradata Vantage on Google Cloud Platform
- Teradata Vantage on Microsoft Azure

Requirements for SAS/ACCESS Interface to Vertica

SAS/ACCESS Interface to Vertica includes SAS Data Connector to Vertica.

SAS/ACCESS Interface to Vertica requires Vertica ODBC Client version 9.1 or later and supports Vertica Analytic Database version 9.1 or later.

To obtain the Vertica Client ODBC driver, contact your database administrator or visit the Vertica website: https://my.vertica.com/download/vertica/client-drivers.

Requirements for SAS/ACCESS Interface to Yellowbrick

SAS/ACCESS Interface to Yellowbrick includes the required PostgreSQL ODBC driver.

Yellowbrick Database version 4.0.0-23452 or later is supported.

Security Requirements

About SAS Viya Security Features

SAS Viya provides Transport Layer Security (TLS) for encryption of data in motion and supports the Advanced Encryption Standard (AES) for encryption of data at rest. With minimal configuration, you can deploy and use the SAS Security Certificate Framework, which contains tools that leverage Kubernetes features to provide encryption for SAS applications.

Multiple methods are supported for user authorization and authentication. Most user account configuration occurs after the deployment has completed. Single sign-on and multi-factor authentication are also supported.

Encryption Overview

The deployment supports TLS version 1.2 or 1.3 for connections to the cluster, from SAS Viya components to your IT infrastructure, and among the Pods. You can deploy in three different modes:

- “Full-Stack TLS”: secure all network connections
- “Front-door" TLS: secure only those components that are intended to accept connections from outside the Kubernetes cluster, such as the Ingress controller and CAS server
- “No TLS”: no network connections are secured. All network transmissions are unencrypted.

In a multi-tenant deployment, all tenants must use the same TLS mode.

To provide TLS, the SAS Security Certificate Framework manages certificates and integrates SAS Viya security with cluster security. The framework supports two options for certificate management. The requirements for each option are described in “TLS Requirements” on page 49.

Deploying in “No TLS” mode is not recommended. The default deployment enables “full TLS,” including TLS for an NGINX Ingress. To secure the Ingress, you can provide your own certificates if desired. Customer-supplied certificates must be in PEM format, and the corresponding private key file is required.

The CAS server supports encryption for tables in caslibs. SAS Viya uses AES with 256-bit keys to encrypt stored data. The encryption applies to source tables, not to tables that are resident in memory. Encryption can be applied to individual tables or to all tables in a library. Each table can

have a unique encryption key, or a single key can be set at the library level in order to have a shared key for all tables in the library.

See the "Encryption Overview" in *Encryption in SAS Viya: Data in Motion* for relevant conceptual information and procedures for customizing your environment.

TLS Requirements

To apply TLS encryption to your SAS Viya deployment, certificates are required. Certificates contain the names of Pods, which are ephemeral. Therefore, a certificate generator that is capable of creating certificates instantly, whenever Pods are scheduled, is required. You can select either cert-manager or a SAS-supplied distribution of OpenSSL as the certificate generator. With either option, the generated CA certificate and private key are stored in a Kubernetes secret in the SAS Viya namespace. This CA certificate and key are used to issue the server identity certificates that secure the SAS Viya back-end servers.

IMPORTANT Additional configuration of your security settings is required. Detailed documentation is provided in the SAS Security Certificate Framework README file. After you have downloaded the deployment assets, the file is located at `$deploy/sas-bases/examples/security/README.md` (for Markdown format) or at `$deploy/sas-bases/docs/configure_network_security_and_encryption_using_sas_security_certificate_framework.htm` (for HTML format).

The following table summarizes the requirements for each option:

Table 17 *Requirements to Support TLS*

Certificate Generator	Requirements
cert-manager	The deployment is configured to use cert-manager by default. If you want to use it, you must deploy cert-manager in your cluster prior to deploying SAS Viya. SAS Viya is compatible with all production releases of cert-manager. Once you have determined the version of Kubernetes that you will be using, check the cert-manager documentation to select a release of cert-manager that is compatible with your version of Kubernetes: https://cert-manager.io/docs/installation/supported-releases/. Do not install more than one instance of cert-manager per cluster. You can download cert-manager from the Jetstack GitHub site. Run the following command to find the currently installed version: `kubectl -n cert-manager describe deployments/cert-manager | grep Image` A required issuer is included by default in the kustomization.yaml file. SAS recommends that you configure cert-manager to automatically delete secrets when they are no longer being used. For instructions, see "Delete Secrets That Are No Longer Used" in *Encryption in SAS Viya: Data in Motion*.

Certificate Generator	Requirements
openssl	The openssl certificate generator is proprietary SAS software that uses the OpenSSL open-source project. It creates a unique CA certificate and private key during SAS Viya deployment. Configure openssl as the SAS Viya certificate generator by following the instructions in the SAS Security Certificate Framework README file.

To support a multi-tenant deployment, the ingress certificate must include a wildcard entry in the SAN DNS attribute in order to match the various tenant names. This requirement applies to the Full-Stack or Front-End TLS modes, regardless of the certificate generator that you select. Without this additional configuration, the browser throws an invalid certificate error as soon as a tenant is onboarded. Multi-tenancy is an optional feature.

Here is an example of a SAN DNS attribute that contains a wildcard: `*.mydomain.mycompany.com`

In an environment where SAS/CONNECT or CAS programming interfaces are used, users might connect directly to a node port or LoadBalancer service from an external IP address or host name, bypassing the ingress controller. If you apply Full-stack TLS to your deployment, additional configuration in the signed certificate that secures the LoadBalancer or NodePort service is required. The certificate's SAN attributes must contain the name or the IP address that is used to connect directly to these services. After you have downloaded the deployment assets for your software order, see "Configuring Certificate Attributes" in the SAS Security Certificate Framework README file.

Pod Security Policies

If your Kubernetes cluster has Pod security policies enabled, a conflict with a SAS Viya seccomp annotation can cause the deployment to fail. SAS Viya requires **`runtime/default`** to be allowed in the environment. You can check Pod security policies by running the following command on your cluster:

```
kubectl get psp restricted -o custom-columns=NAME:.metadata.name,
"SECCOMP":".metadata.annotations.seccomp\.security\.alpha\.kubernetes\.io/
allowedProfileNames"
```

You can resolve this conflict by updating a variable in the pod security policy. For more information, see the following SAS Note: https://support.sas.com/kb/67/349.html.

DNS Requirements for Multi-Tenancy

As the users within each tenant access SAS Viya components, the host name that they use to access the SAS deployment identifies their tenant membership. Therefore, DNS records for tenant-specific subdomains are required.

Each tenant is reachable by a URL that is derived from the provider's URL. Here is the format for a typical tenant URL

`tenant-ID.provider-ingress`

Here is an example of a provider URL:

```
sasviya.mycompany.com
```

Here is an example of a tenant URL:

```
mytenant.sasviya.mycompany.com
```

You must verify that the DNS server for your enterprise is configured to route to these address spaces. You can create a wildcard subdomain entry as a time-saving step.

Identity Provider and Authentication Requirements

SAS Viya supports LDAP and the System for Cross-domain Identity Management (SCIM) for user and group identities. Multiple identity providers are supported, including Microsoft Azure Active Directory.

Supported Authentication Methods

SAS Viya supports the following methods for authenticating users who are signing in to the environment:

- LDAP, Kerberos, and single sign-on
- Single sign-on using Security Assertion Markup Language (SAML) or OpenID Connect

  Single sign-on support is limited to web log-ins.
- A pluggable authentication module (PAM) can be used to validate the user's credentials when accessing the CAS server.

  PAM is supported only for CAS connections, not for web log-ins. This type of authentication enables support for users to launch CAS server sessions under their host identities.

  PAM uses the operating system for user and password authentication, which means that you must also set up System Security Services Daemon (SSSD) and enable host authentication. Similar requirements apply to Kerberos.

Host Authentication

Users launch CAS sessions under a shared identity by default. For various reasons, you might instead want to enable them to launch sessions under their host (operating-system) accounts. With host identities, users launch CAS and compute server sessions under their own user accounts, defined in the operating system. Configure host authentication in order to enable Kerberos authentication, to facilitate access to NFS volumes, to enable some users to deploy models written in Python or R, or to integrate with a SAS®9 deployment.

The ability to launch CAS sessions under a host identity is disabled by default. To enable it, you must apply an overlay to the base kustomization.yaml file and perform additional configuration when the deployment has completed. For more information, see "Enable Host Launch" in *SAS Viya: Deployment Guide*.

LDAP Requirements

When the deployment completes, SAS Viya is configured by default to use LDAP. The following requirements apply to your LDAP server:

- SAS Viya must have Read access to your LDAP server.

- In order to bind to the LDAP server, SAS Viya requires either a system account (with a userDN and password) or anonymous binding.
- If the mail attribute is specified for LDAP accounts, it must have a non-null value that is unique for each user.
- LDAPS is supported, but the required certificates are not configured automatically by the deployment process.

You will configure SAS Viya identities as a post-deployment step. Instructions for setting up identities with LDAP are provided in Identity Management.

Additional LDAP Requirements for Multi-Tenancy

You cannot use multiple LDAP servers for a single tenant. However, you have the following options:

- One LDAP server per tenant

  You can specify custom LDAP properties for each tenant.

- A single LDAP server for all tenants

  You can specify custom LDAP properties for each tenant.

  A separate OU per tenant and an OU named “provider” are required if you also use the same LDAP directory structure for all tenants.

You should either set up or plan your tenant structure in LDAP before you start the deployment. Determine whether you will use the same directory structure for the users and groups within all tenants (a “fixed” LDAP structure), or will use a custom structure that varies per tenant. Based on these decisions, you can then perform tenant onboarding as a post-deployment task.

Multi-tenancy requires some post-deployment configuration in SAS Environment Manager. By default, the values that you specify for tenant LDAP connection parameters are automatically applied to the provider and to the users and groups within all tenants. However, you will have an option to selectively apply LDAP connection settings. This option enables you to deploy a custom directory structure for each tenant. For example, you can use a single LDAP server across all tenants while using custom parameters, such as the baseDNs or search filters, for each tenant. In such a case, you would select the option to **Apply configuration only to this tenant (provider)** for the provider’s group and user connections.

During tenant onboarding, if you select the option to apply the configuration only to the provider, you must use the fixed directory structure that is described in “User Accounts for Multi-Tenant Deployments: Single LDAP Server for All Tenants” on page 58. The reason is that SAS Viya requires an OU for the provider and separate OUs for each tenant if the option is not selected. For more information about your options for setting up tenants in LDAP, see “User Accounts for Multi-Tenant Deployments: Separate LDAP Server per Tenant” on page 59.

SCIM Requirements

SAS Viya supports SCIM 2.0.

When the deployment completes, SAS Viya is configured by default to use LDAP. If you intend to use SCIM, you must disable LDAP by logging in to SAS Environment Manager after the deployment process has completed.

Single sign-on using SAML or OpenID Connect (OIDC) is required if you configure a SCIM identity provider.

With SCIM, you can use Okta or Microsoft Azure Active Directory to populate user and group identities in SAS Viya. In a Google Cloud Platform (GCP) environment, Google Identity is not yet supported for user provisioning; use Okta or Azure Active Directory instead. SCIM requires network access to the SAS Viya deployment using HTTPS, and it requires a certificate signed by a public certificate authority. The SCIM IdP also requires a long-lived token to access the SAS Viya APIs.

Check your firewall settings so that IP addresses used by the SCIM IdP are allowed to reach the SAS Viya Ingress. In Microsoft Azure, you can configure your network security group to allow Azure Active Directory to communicate with resources in your virtual network by enabling inbound access for the AzureActiveDirectory service tag.

If you use OIDC, SAS Logon Manager must construct the correct redirect_uri parameter to send as part of the authentication request to the IdP. This parameter incorporates header values from the request to SAS Logon Manager. The relevant values are from the Host, X-Forwarded-Proto, and X-Forwarded-Port headers. The redirect_uri parameter must match the value or values that are registered with the OIDC provider. Its value is also used by the OIDC provider to redirect the client browser back to SAS Logon Manager after authentication with the authorization code. Therefore, these header values must correspond to the external address that is used by SAS Viya users.

In an OIDC environment with a reverse proxy server in front of an NGINX Ingress Controller, the NGINX configuration setting `use-forwarded-headers` must be changed from the default "false" to "true". Changing this setting enables the ingress controller to pass the incoming X-Forwarded-* headers from your reverse proxy to SAS Logon Manager. SAS Logon Manager can then build the redirect_uri parameter correctly based on those headers.

You configure SAS Viya identities as a post-deployment step. Instructions for setting up identities with SCIM are provided in Identity Management.

Additional SCIM Requirements for Multi-Tenancy

SCIM configuration for a deployment with multi-tenancy occurs after the SAS Viya deployment has completed. The first step is to onboard tenants. You can then register an OAuth client on each tenant. Although you can typically register a client manually or programmatically, using the SAS Viya Command-Line Interface, a multi-tenant deployment only supports a manual client registration. See “Register a New Client ID” in *SAS Viya: Authentication* for the steps.

SCIM setup procedures are described in detail in “Configure SCIM Provisioning in the Identity Provider ” in *SAS Viya: Identity Management*. The configuration steps must be repeated for each tenant.

Kerberos Requirements

Configure Kerberos delegation before you start the deployment. If you plan to use SAS/ACCESS Interface to Hadoop to connect to a Hadoop data source, only unconstrained delegation is supported.

Enabling Kerberos in your deployment also requires that you enable host authentication. By default, users connect to the CAS server and launch sessions under a shared service account. However, use of the shared account is not supported in a Kerberos environment.

You have two options for enabling host authentication for Kerberos. Both options require post-deployment steps. For more information, see “Configure Kerberos” in *SAS Viya: Authentication*.

Requirements for User Accounts and Services

About Roles and Permissions

The Kubernetes specification accepts system-delineated roles and user-facing roles. SAS Viya applies a *least-privileges* model that grants to each role the minimum access that is required to deploy and run the application. SAS Viya resources are controlled by labels that describe installation requirements and that determine the privileges for enabling access to these resources.

To deploy SAS Viya, an administrator or administrators with sufficient permissions must run the commands to apply the application resources to the cluster. The requirement to use label selectors applies to kubectl commands that require the highest level of access: `kubectl apply` and `kubectl delete`. For `kubectl apply`, the label selector explicitly avoids race conditions. Typical race conditions occur when the Kubernetes API processes objects in parallel that have dependencies on each other.

Here is an example of a kubectl command that uses a label selector:

```
kustomize build | kubectl apply --kubeconfig=cluster-admin.conf --selector="sas.com/admin=cluster-wide" -f -
```

The selector `sas.com/admin=cluster-wide` indicates that the command can be executed only by a user with the default cluster-admin ClusterRole. This role grants the equivalent of super-user privileges.

To provide granular privileges and to ensure that the least-privileges model is enforced, SAS applies a custom resource label to one category of SAS Viya resources. The following table summarizes SAS Viya resources and their corresponding labels:

Table 18 *SAS Viya Label Selectors*

Label	Description	Examples of SAS Viya Resources
cluster-wide	Applies to resource modifications that affect the entire cluster. This category includes global objects, such as CRDs or ClusterRoles, that could affect deployments in other namespaces. Interacting with resources in this category also requires cluster-admin permissions. These resources are defined under Cluster APIs in the Kubernetes API Reference documentation.	CustomResourceDefinition; ServiceAccount; ClusterRole; Role; PersistentVolume
cluster-local	A SAS custom label for resources that require cluster-admin permissions, but it limits the impact of changes to the	ResourceQuota; LimitRange; Secret; ConfigMap; RoleBinding;

Label	Description	Examples of SAS Viya Resources
	namespace, with a reduced impact on the cluster. These resources are defined under Config and Storage APIs in the Kubernetes API Reference.	ClusterRoleBinding; PersistentVolumeClaim; PodTemplate
cluster-api	A Kubernetes label for cluster-wide resources that are specific to CustomResourceDefinitions. This includes the CRDs themselves as well as any supporting conversion webhook resources. As CRDs extend the API supported by the cluster where they are created, they might require additional consideration.	CustomResourceDefinition; Deployment; Service
namespace	Applies to the following resource types: ■ resources for managing and running containers ■ resources that provide external access to workloads by means of a LoadBalancer or NodePort service These resources are defined under Workloads APIs, Service APIs, and Metadata APIs in the Kubernetes API Reference.	Service; Deployment; ReplicaSet; ReplicationController; Job; CronJob; StatefulSet; DaemonSet

Note: A Kubernetes user with sufficient permissions to create these SAS Viya resources must also have permissions to get, list, watch, update, patch, and delete these resources.

Some Kubernetes Role-Based Access Control (RBAC) policies are enabled by default. Because objects of each kind are deployed, permissions for each resource should be granted to the appropriate role. During RBAC planning for your SAS Viya deployment, consider that some user-facing ClusterRoles need to allow admin users to include rules for custom resources.

Cluster Resources and Roles That Require Elevated Permissions

As a cluster administrator, you should understand the full set of requirements for administrators and have some knowledge of the cluster-wide resources that are used by SAS Viya. Cluster administrators can assign permissions to other users, such as namespace administrators. These delegated administrators might perform some deployment steps, or they might need permissions to view pod-level or cluster-level information after the deployment has completed.

Depending on the security architecture that is in use, the cluster administrator must deploy certain cluster-level resources as part of a SAS Viya deployment. These resources might include custom resource definitions (CRDs), Roles, RoleBindings, and PodTemplates. Once the resources have been

deployed, SAS recommends providing a delegated, namespace-level administrator with the "get", "list," and "watch" permissions on all resources cluster-wide.

With these permissions, the delegated administrator has broad Read-Only access to relevant cluster-level resources. In order to effectively monitor the SAS workload, the namespace-level administrator might also be granted permissions to view cluster metrics (for example, by using Lens, an integrated development environment for Kubernetes). These Read-Only permissions might range more broadly than absolutely required, but they provide the delegated administrator with an enhanced ability to monitor the cluster and make recommendations to the cluster administrator.

A SAS Viya deployment adds custom API extensions (“API groups”) to the cluster. The delegated administrator should receive full permissions (all verbs, cluster-wide) for these custom API groups. The following table provides a list of the custom API groups and CRDs that SAS provides:

Table 19 *Custom SAS Viya API Groups*

API Group	CRD	Purpose
opendistro.sas.com	OpenDistroCluster	Used to deploy the Open Distro for ElasticSearch database.
(For only SAS Event Stream Processing or SAS Analytics for IoT) iot.sas.com	ESPConfig	Used to configure the ESP operator and its custom resources.
	ESPLoadbalancer	Used to manage the load balancers for SAS Event Stream Processing.
	ESPRouter	Used to manage the process of connecting SAS Event Stream Processing sources to destinations.
	ESPServer	Used to manage ESP servers.
	ESPUpdate	Used to manage operations across multiple ESP servers.
viya.sas.com	CASDeployment	Used to deploy the CAS server.
webinfdsvr.sas.com	Pgclusters	Stores information that is required to manage a PostgreSQL cluster.
(For an internal PostgreSQL server only) crunchydata.com	Pgpolicies	Stores a reference to an SQL file that can be executed against a PostgreSQL cluster. Crunchy Data provides the internal instance of SAS Infrastructure Data Server.

API Group	CRD	Purpose
	Pgreplicas	Stores information that is required to manage the replicas within a PostgreSQL cluster.
	Pgtasks	A general-purpose CRD that accepts a type of task that is needed to run against a cluster. For information about how the Crunchy Data PostgreSQL Operator uses custom resources (CRs), see: https://access.crunchydata.com/documentation/postgres-operator/4.7.0/custom-resources/.
(Optional) orchestration.sas.com	SASDeployment	Used when you deploy SAS Viya by using the SAS Deployment Operator.

The delegated administrator might also need Write access to a minimal set of resources in the SAS Viya namespace. The following namespace-level permissions enable the administrator to execute a command in a container within a pod. For example, the administrator can inspect processes, check user and group IDs, verify that persistent volumes have been mounted and get their paths, and more:

```
rules:
- apiGroups: [""]
  resources: ["pods/exec"]
  verbs: ["create"]
```

IMPORTANT If you use the SAS Deployment Operator for the deployment, the SASDeployment CR is also deployed before you begin the actual SAS Viya deployment process. The CRD is associated with the orchestration.sas.com api group. The SAS Deployment Operator requires cluster-admin privileges in order to create the CR. After it is deployed, the operator runs with cluster-admin privileges.

If your security architecture allows namespace-level administrators to create role bindings, a few role bindings require specific RBAC permissions. These permissions are in addition to those that are already granted to namespace-level accounts that have the default “admin” or “edit” cluster roles. The following table summarizes these permission requirements and provides examples of roles, which vary based on the offerings in your order:

Table 20 *SAS Viya Custom Permissions by Role*

Examples of Roles	Custom Permissions	Description
sas-data-server-operator-leader-election-role-binding sas-opendistro-operator	Permissions to "delete", "patch", or "update" events	Roles that manage SAS Viya components (SAS Infrastructure Data Server, SAS Open Distro for

Examples of Roles	Custom Permissions	Description
sas-opendistro-operator-leader-election		Elasticsearch, and SAS Cache Server).
(For only Red Hat OpenShift) sas-cas-operator	Permissions to specific API groups: monitoring.coreos.com, route.openshift.io, and projectcontour.io	
sas-data-server-operator	Permissions to specific API group: coordination.k8s.io	

User Accounts

You can set up SAS Viya users and groups with SAS Environment Manager after your software has been deployed.

Some pods run system-critical processes under the UID 1001. This UID cannot be changed. Verify that no user accounts in your LDAP directory are using this UID.

An account that is named sasboot is also created during the deployment and has a password that will expire when used. It is an administrator account that is used for preliminary logon to SAS Environment Manager. It is internal only to SAS and does not exist in your identity provider. For more information, see Sign In as the sasboot User.

User Account Requirements for Multi-Tenant Deployments

Requirements for multi-tenancy depend on whether you intend to use a single LDAP server for all tenants, or one LDAP server per tenant. Prepare the LDAP server before you start a deployment with multi-tenancy enabled.

User Accounts for Multi-Tenant Deployments: Single LDAP Server for All Tenants

The LDAP directory structure that is described here is required if you decide to apply the same connection settings to the provider and to users and groups within all tenants during tenant onboarding (a post-deployment task). These steps are recommended for an environment in which a single LDAP server is used for all tenants. You can also use a custom structure that applies different settings to different tenants, and you can use a separate LDAP server per tenant. For more information, see “User Accounts for Multi-Tenant Deployments: Separate LDAP Server per Tenant” on page 59.

To configure the LDAP server:

1 Create the provider OU. Here is an example:

```
dc=example,dc=com
       ou=tenant-1
           ou=groups
           ou=users
       ou=tenant-2
           ou=groups
           ou=users
       ...
       ou=provider
           ou=groups
           ou=users
```

Here is an example that uses LDIF syntax:

```
dn: cn=sas,ou=groups,ou=provider,dc=sas,dc=com
distinguishedName: cn=sas,ou=groups,ou=provider,dc=sas,dc=com
displayName: Tenant-admin-group-for-provider
gidNumber: value
objectClass: groupOfUniqueNames
objectClass: extensibleObject
uniqueMember: uid=sas,ou=people,ou=provider,dc=sas,dc=com
cn: sas
```

Note: The provider DN must be specified as **`provider`**.

2 For each tenant user that you define in LDAP, the following requirements apply:
   - Each uidNumber with gidNumber attributes has been specified.
   - Each user ID, across all tenants, is unique.
   - The homeDirectory attribute is set to a value that is appropriate for your environment.
   - The loginShell attribute is set to /bin/bash.

3 (Optional) If your provider or tenants will have secondary CAS controllers to enable failover, set up a shared file system.

The deployment process automatically creates an internal user account for an administrator within the provider tenant. You can set up separate groups for administrative users and for non-administrative users within each tenant in LDAP, and you can add tenant users to one of these groups. The tenant creation process provides these groups with access to critical files and other resources that are otherwise restricted. The users that are defined within tenants should not be added to overarching administrators' groups. Tenant users must instead be members of their tenant's user groups.

Tenant onboarding is a post-deployment task. With the directory structure that is described here, do not select the option to **Apply configuration only to this tenant (provider)** when you perform tenant onboarding. The settings are then applied automatically to all tenants. Any tenants that you add in the future will require similar directory settings.

User Accounts for Multi-Tenant Deployments: Separate LDAP Server per Tenant

SAS Viya supports an environment in which a separate LDAP server is used for each tenant. You can customize the LDAP directory structure or use your existing structure. To prepare your environment for tenant onboarding:

1. Before you update the kustomization.yaml file, plan and document the structure of the tenant spaces in the LDAP servers that you will allocate to tenants.
2. For each tenant user that you define in LDAP, the following requirements apply:
   - Each uidNumber with gidNumber attributes has been specified.
   - Each user ID, across all tenants, is unique.
   - The homeDirectory attribute is set to a value that is appropriate for your environment.
   - The loginShell attribute is set to /bin/bash.
3. Once the deployment process has completed, log on to SAS Environment Manager using the sasboot account.
4. Select settings for the following parameters to enable the Identities service to authorize the provider:

   ```
   sas.identities.providers.ldap.connection
   sas.identities.providers.ldap.group
   sas.identities.providers.ldap.user
   ```

   As you configure each parameter, consider whether to select the option to **Apply configuration only to this tenant (provider)**. This option restricts the application of a setting to the provider. If you do not select this option, the values for each parameter are applied to all tenants, and you must set up the directory structure as specified in “User Accounts for Multi-Tenant Deployments: Single LDAP Server for All Tenants” on page 58.
5. (Optional) If your provider or tenants will have secondary CAS controllers to enable failover, set up a shared file system.

You can set up separate groups for administrative users and for non-administrative users within each tenant in LDAP, and you can add tenant users to one of these groups. The tenant onboarding process provides these groups with access to critical files and other resources that are otherwise restricted. Take care to ensure that tenant users are members only of their tenant’s user group.

Service Accounts

During the deployment process, required infrastructure services are automatically started in separate containers within the SAS Viya namespace. Some of these services are owned by a Kubernetes service account, which is created by the deployment process. Each service account has a very limited role. The use of multiple service accounts to run services makes it easier to secure your SAS Viya environment because each service account has only the specific privileges that it requires to run one service and has no additional privileges.

The following Kubernetes service accounts are created by the deployment process:

Table 21 *Required Kubernetes Service Accounts*

Service Account	Description
postgres-operator	Owner of the PostgreSQL operator. Supports either an internal or an external instance of PostgreSQL for SAS Infrastructure Data Server.

Service Account	Description
pgo-backrest	Owner of the Crunchy Data service to support an internal PostgreSQL instance for SAS Infrastructure Data Server. All the containers that are launched by the PostgreSQL operator as part of the pgbackup service use this service account. Crunchy Data service role bindings are part of the custom resource definition for the PostgreSQL operator. This account and container are not present if you selected to deploy with an external instance of PostgreSQL.
pgo-default	Used by the sas-crunchy-data-postgres-backrest-shared-repo service.
pgo-pg	Used by the Crunchy Data service, sas-crunchy-data-postgres.
pgo-target	Owner of the Crunchy Data service account to support an internal PostgreSQL instance for SAS Infrastructure Data Server. Crunchy Data service role bindings are part of the custom resource definition for the PostgreSQL operator. This account and container are not present if you chose to deploy with an external instance of PostgreSQL.
sas-rabbitmq-server	Owner of the SAS Message Broker service.
sas-cas-operator	Owner of the CAS server operator.
sas-cas-server	Owner of the CAS server pods. It enables CAS server pod security policies. This account has both a role and a role binding of sas-cas-server.
sas-config-reconciler	Owner of the service that keeps pods synchronized with configuration data, relaunching them when their configuration changes.
sas-consul-server	Owner of the Consul key-value store. SAS Configuration Server is based on HashiCorp Consul.
sas-data-server-utility	Owner of a required account to support SAS Infrastructure Data Server. Manages database creation, Consul registration, and PostgreSQL validation. It runs in a container that is separate from the PostgreSQL Operator. This container is always present, regardless of whether you deployed with an internal or external PostgreSQL instance.
sas-launcher	Owner of a service that launches Kubernetes jobs and checks the status of those jobs.
sas-prepull	A service account that improves start-up times by pre-staging the sas-programming image to nodes that require it.

Service Account	Description
sas-programming-environment	Supports a group of components that include the compute server, the batch server, and SAS/Connect.
sas-viya-backuprunner	A service account that is used for backup operations.

Additional service accounts might be present if you deployed SAS Viya products that require them.

To run on Red Hat OpenShift, some of these service accounts require a security context constraint (SCC) statement. For more information, see “Requirements for Security on Red Hat OpenShift” on page 62.

Requirements for Security on Red Hat OpenShift

If you are deploying SAS Viya on Red Hat OpenShift, a few additional steps are required to prepare the cluster for the SAS Viya deployment. If you are not deploying on OpenShift, you can skip this section.

SCCs and Pod Service Accounts

In a Red Hat OpenShift environment, each Kubernetes Pod is started with an association with the restricted security context constraint (SCC) provided by Red Hat, which limits the privileges that each Pod can request.

Most SAS Viya Pods are deployed in the restricted SCC, which applies the highest level of security. Two other OpenShift predefined SCCs are used by default. In addition, a few custom SCCs are either required by essential SAS Viya components, such as the CAS server, or associated with specific SAS Viya offerings that might be included in your software order.

For a full list and description of the required SCCs, see “Security Context Constraints and Service Accounts” in *SAS Viya: Deployment Guide* .

SCCs and File System Permissions

SAS Viya includes default fsGroup settings that enable file system access. When an fsGroup ID is set for a Pod, any files that are written to a volume within that Pod inherit that fsGroup as their group ID (GID). The fsGroup ID is the owner of the volume and of any files in that volume.

For additional security, most Pods are set to use an fsGroup value (1001) that is not supported by the OpenShift restricted SCC. These values must be modified before you start the deployment process.

The following table summarizes the default fsGroup settings:

Table 22 *Default fsGroup Values*

SAS Viya Component	Default fsGroup Value	Explanation
CAS server	1001	Changing this value is optional because a custom SCC definition enables the shared service account to access the CAS server. The `$deploy/sas-bases/examples/cas/configure/cas-server-scc.yaml` file grants SCCs for the service account GID by default. If you plan to enable users to launch CAS sessions under their host identities, use cas-server-scc-host-launch.yaml instead in order to set the correct capabilities and privilege escalation. For more information, see "Security Context Constraints and Service Accounts" in *SAS Viya: Deployment Guide* .
PostgreSQL Server	1001 (postgres service) 2000 (pgbackrest service)	These values are required for the internal PostgreSQL server (SAS Infrastructure Data Server) and cannot be modified. Custom SCC definitions enable their usage. For more information, see "Security Context Constraints and Service Accounts" in *SAS Viya: Deployment Guide* .
Open Distro for Elasticsearch	1000	The Open Distro for Elasticsearch components are assigned to a custom SCC with the same restrictions on the fsGroup as the OpenShift restricted SCC. The default fsGroup setting does not enable deployment and must be modified in a pre-deployment step.
All other Pods	1001	All remaining Pods are assigned to the OpenShift restricted SCC. This SCC does not enable containers to use the default fsGroup settings and must be modified in a pre-deployment step.

SAS provides the update-fsgroup.yaml file to help you update the fsGroup in targeted manifests with the correct GID value. The required steps to change the fsGroup setting are provided in the Container Security README. For more information, see Additional Security for Red Hat OpenShift.

If you use the optional section in update-fsgroup.yaml to update the fsGroup for CAS pods, make sure that you also update the fsGroup range in the SCC that is applied to the CAS service account by using one of the provided YAML files that are specific to CAS.

Removing the seccomp Profile

Most Pods run under the restricted SCC. However, some settings cannot be included in the podSpec if it runs under the restricted SCC, such as a seccomp profile. These settings must be removed for a deployment on OpenShift.

The Container Security README provides instructions for removing the seccomp profile using customizations provided by SAS. For more information, see Additional Security for Red Hat OpenShift.

PostgreSQL Deployment Options

SAS Infrastructure Data Server stores SAS Viya user content, such as reports, authorization rules, selected source definitions, attachments, audit records, and user preferences. SAS Viya uses High Availability (HA) PostgreSQL for SAS Infrastructure Data Server. Take some time to consider your options for deploying a PostgreSQL instance to support the data server.

Internal versus External PostgreSQL Instances

An instance of PostgreSQL is required for the SAS Data Infrastructure Server. You can allow SAS to automatically deploy an instance of PostgreSQL, which is referred to as an *internal* instance, or you can provide your own PostgreSQL, which is an *external* instance. Before deploying your software, you must decide which type of instance you want.

IMPORTANT After your SAS Viya software has been deployed, you cannot change the PostgreSQL deployment type without redeploying the entire software package.

In addition to the following requirements, both types of PostgreSQL instance require persistent storage and PVCs. For more information, see “Persistent Storage Volumes, PersistentVolumeClaims, and Storage Classes” on page 15.

Note: SAS Model Risk Management does not support an external PostgreSQL database instance.

Internal PostgreSQL

If you decide to deploy an internal instance of PostgreSQL for SAS Infrastructure Data Server, SAS configures and maintains the deployment for you.

SAS deploys a comprehensive PostgreSQL container operator suite provided by Crunchy Data for this purpose.

As a result of the recent change in supported versions of Kubernetes, SAS Viya with an internal PostgreSQL instance must be fully updated to at least version 2021.2.3 before you can update any additional deployments to version 2021.2.4. A Kubernetes cluster that contains multiple deployments of SAS Viya with internal instances of PostgreSQL is a supported topology. However, multiple deployments that run in separate namespaces might share deployment resources. If your deployments are not updated to the same version of SAS Viya, the different Crunchy Data instances might create compatibility issues.

Additional Requirements for Red Hat OpenShift

If you are not deploying SAS Viya on OpenShift, you can skip this section.

Security Context Constraints for fsgroup 26 and fsgroup 2000 are required for the internal PostgreSQL server. A Kubernetes cluster administrator must add these SCCs to their OpenShift cluster prior to deploying SAS Viya.

After mounting the required PVC, the internal PostgreSQL server cluster changes the ownership of all the files and directories to the predefined PostgreSQL user ID and group ID of 26. An additional Pod that handles and stores backups of the SAS Infrastructure Data Server changes the ownership of files and directories to a user ID and group ID of 2000. The cluster is therefore unable to start up unless permission for the required SCCs is granted.

After you have downloaded and uncompressed the deployment assets, the `$deploy/sas-bases/examples/crunchydata/openshift` directory contains a file to grant SCCs for fsgroup 26 and 2000 on an OpenShift cluster. For more information, see “Security Context Constraints and Service Accounts” in *SAS Viya: Deployment Guide* .

Requirements for External PostgreSQL

If you decide to deploy an external instance of PostgreSQL, you are responsible for configuring and maintaining the PostgreSQL deployment. Providing your own instance enables you to restrict the ability of SAS services to access the database when a single PostgreSQL instance houses multiple databases. The following requirements apply:

- PostgreSQL versions 11 and 12 are the only supported versions of PostgreSQL for SAS Viya.
- Managed PostgreSQL instances that are hosted by your cloud provider, such as Microsoft Azure Database for PostgreSQL, Amazon RDS for PostgreSQL, or Cloud SQL for PostgreSQL, are supported.
- All PostgreSQL extensions must be installed.
- The external PostgreSQL instance must already be running and the database owner must be created before you start the SAS Viya deployment.
- Depending on the database that you provided, the database or schema owner requires permissions to Create, Read, Update, and Delete (CRUD). For example, if the PostgreSQL instance that you provide does not yet have the SAS Viya database created, the deployment process creates it for you, using the database owner that you provided. Adjust user roles, database permissions, and attributes accordingly.
- To enable some SAS Viya features, such as backup, restore, and multi-tenancy, the user ID of the database owner should be created with CREATE ROLE and CREATEDB permissions. In addition, this user must be the owner of the initial database that is created for SAS Viya in the external PostgreSQL instance.

  Refer to the following PostgreSQL documentation for details: https://www.postgresql.org/docs/12/ddl-priv.html.
- A low-latency, high-bandwidth environment is required.

  Placing the external PostgreSQL database in a separate data center, region, or availability zone from the rest of your SAS Viya components might lead to increased latency and reduced

bandwidth. Such conditions are likely to cause a degraded overall performance of the environment. Before attempting to deploy SAS Viya with an external PostgreSQL database, it is important to test with it in order to balance cost and performance considerations and to confirm that performance is not adversely affected.

- You must provide the CA certificate that is used to secure the PostgreSQL database server to your SAS Viya deployment.

  As part of securing the database server, your cloud provider might provide a CA certificate. You must make this certificate available during the deployment process so that it can be imported into the SAS Viya certificate infrastructure. For more information about how certificates are managed, see the “Incorporating Additional CA Certificates into the SAS Viya Deployment” section of the SAS Security Framework README file, located at **`$deploy/sas-bases/examples/security/README.md`** (for Markdown format) or at **`$deploy/sas-bases/docs/configure_network_security_and_encryption_using_sas_security_certificate_framework.htm`** (for HTML).

- (GCP only) You must use Google’s Cloud SQL Proxy to access your Cloud SQL for PostgreSQL server.

  Additional steps are required in order to support full TLS. For more information, see the README file at **`$deploy/sas-bases/examples/postgres/configure/README.md`** (for Markdown format) or at **`$deploy/sas-bases/docs/configure_postgresql.htm`** (for HTML).

PostgreSQL Requirements for a Multi-Tenant Deployment

SAS Infrastructure Data Server is an important component for multi-tenancy and provides data isolation. You have two options for setting up this isolation:

- a separate PostgreSQL database for each tenant

  A separate database instance is provisioned for each tenant. This option is recommended if maximum isolation of tenant data is required.

  Unique database credentials per tenant enhance security and isolation. However, each tenant must have a unique database connection pool, which might significantly increase the total connection count that the back-end database server must support. Additional tuning is required.

- a shared PostgreSQL database that is partitioned to provide tenant isolation

  A single database is shared by all tenants, but each tenant is partitioned into separate schemas based on the SAS Viya service. The schema's name is generated based on the application and tenant name (such as identities_tenantA).

  This option is recommended when database connection resources are limited. A single connection pool is used for all tenants. However, data for all tenants is secured by a single credential. In addition, because connections for all tenants come from a single connection pool, a single tenant can consume all connection resources, depriving other tenants.

Mixing internal and external PostgreSQL database instances among tenants is not supported.

For either option (separate databases per tenant or separate schemas per tenant), you should configure additional connections on the database server. Spikes in connection usage have been observed during tenant onboarding and when users log in and start using SAS Viya. The baseline recommendation for SAS Viya is to set `max_connections` to a minimum of 1024 for an external data server. The internal data server is set to 1280 max connections by default. However, this setting is partially dependent on the SAS Viya offerings that you have purchased.

For an external or internal PostgreSQL data server, you can use the following baseline formula to size your environment and tune the settings to improve performance:

```
(number of tenants + 1) * 1128 = max_connections
```

For example, if you plan for three tenants, (3 tenants + 1 provider) * 1128 = 4512 max_connections.

In order to prepare for peak usage during tenant onboarding and SAS Viya system usage, temporarily allocate additional connections:

- 20% more connections for the database-per-tenant option
- 25% more connections for the schema-per-tenant option

Note: Additional PostgreSQL tuning might be required after the deployment process has completed.

You can use a ConfigMap to change the default value for max_connections. After you have downloaded the deployment assets, an example file for an internal PostgreSQL server, sas-postgres-custom-config.yaml, is provided in your **`$deploy/sas-bases/examples/postgres/custom-config`** directory. For more information about tuning an internal PostgreSQL server, see "Updating the Configuration of a PostgreSQL Cluster" in *SAS Viya: Infrastructure Servers*.

Open Distro for Elasticsearch Requirements

SAS Viya includes Open Distro for Elasticsearch, which is an Apache 2.0-licensed distribution of Elasticsearch with enhanced security. SAS Viya uses it to support search features.

Additions to the kustomization.yaml file are required in order to configure Open Distro for Elasticsearch. For more information, see Configure Open Distro for Elasticsearch.

Modify Default Virtual Memory Resources

The Open Distro for Elasticsearch Pods require additional virtual memory resources. All nodes that run workloads in the stateful workload class are affected by this requirement. In order to provide these memory resources, a transformer can use a privileged container to set the virtual memory for the mmapfs directory to the required level. However, privileged containers must be permitted by your Pod security policies.

If privileged containers are enabled, you can add a transformer, sysctl-transformer.yaml, to the base kustomization.yaml file and configure the corresponding overlay. Otherwise, you have other options for configuring memory resources before you start the deployment process, including modifying settings manually on each node.

If you are using a managed Kubernetes cluster, your cloud provider probably provisions the nodes dynamically. In this instance, be aware that manual modifications do not persist after a restart of a Kubernetes node. The cluster administrator must use an alternative method to save the vm.max_map_count setting.

For information about all of your options for managing memory settings, see Configure Default Virtual Memory Resources.

Provision Storage

Open Distro for Elasticsearch requires a StorageClass that provides persistent block storage or a local file system mount in order to store the search indices. Remote file systems, such as NFS, are not supported for this purpose. However, the default storage class is typically adequate. Here are some examples of StorageClass options that meet the minimum requirements:

Table 23 *Storage Options per Platform*

Platform	StorageClass
Microsoft Azure	default (kubernetes.io/azure-disk)
Amazon Web Services	gp2 (kubernetes.io/aws-ebs)
Google Cloud Platform	standard (kubernetes.io/gce-pd)
Red Hat OpenShift	thin (kubernetes.io/vsphere-volume)

Note: In a multi-tenant deployment, all tenants use the same Open Distro for Elasticsearch resources.

A minimum of one PVC is required, with accessMode RWO. The PVC is typically created automatically as part of the deployment, with a default size of 128 Gi.

To help you customize your deployment to apply the required StorageClass and transformer, an example file for Open Distro for Elasticsearch has been provided in `$deploy/sas-bases/examples/configure-elasticsearch/internal/storage/`. The README file in the same directory provides instructions, or see `$deploy/sas-bases/docs/configure_a_default_storageclass_for_open_distro_for_elasticsearch.htm` for the instructions in HTML format.

For more information about storage requirements, see “Persistent Storage Volumes, PersistentVolumeClaims, and Storage Classes”.

Configure a Storage Class for Red Hat OpenShift

Open Distro for Elasticsearch requires a StorageClass that provides persistent storage for the search indices. For OpenShift in a vSphere private cloud, the thin StorageClass is an example of a StorageClass that is appropriate for Open Distro for Elasticsearch. However, you must customize your deployment to point to the required StorageClass and transformer.

The example YAML file that is shown in VMware vSphere object definition creates an appropriate storage class on the specified VMware data store. The SAS administrator can then use that data store for the VMware VMDK files.

To help you customize your deployment to apply the required StorageClass and transformer, an example file for Open Distro for Elasticsearch has been provided in `$deploy/sas-bases/examples/`

`configure-elasticsearch/internal/storage`. The README file in the same directory provides instructions, or see `$deploy/sas-bases/docs/configure_a_default_storageclass_for_open_distro_for_elasticsearch.htm` for the instructions in HTML format.

On OpenShift, you must also configure permissions in VMware vSphere to enable the provisioning of the storage option that you select. The user that is specified in the install-config.yaml file for the vSphere installation must have the roles and privileges that are required for persistent volume provisioning. The required permissions depend on whether you provision static or dynamic storage.

SAS recommends that you use an OpenShift plug-in for vSphere to enable the StorageClass. The OpenShift documentation describes multiple storage options that SAS Viya supports. You can follow the instructions that Red Hat provides in Persistent storage using VMware vSphere volumes to set up persistent storage for Open Distro for Elasticsearch.

Additional Configuration for OpenShift

Open Distro for Elasticsearch on OpenShift requires that you apply some Security Context Constraints. Deploying on OpenShift also requires changes to a few kernel settings. For more information, "Security Context Constraints and Service Accounts" in *SAS Viya: Deployment Guide*.

Client Requirements

Web Browsers

End users can access the product user interfaces for SAS Viya applications from a desktop computer, using a supported web browser. Because SAS software is not installed on this machine, the requirements are minimal. UNIX and 64-bit Windows operating systems are supported.

Your browser must be enabled for JavaScript.

Some SAS Viya user interfaces include some advanced features that require recent versions of popular web browsers. Client computers that access these interfaces should adhere to the following guidelines:

Table 24 *Supported client operating systems and browsers*

Operating System	Web Browser (64-bit)	Web Browser (32-bit)
Apple macOS 10.13 and later	Apple Safari 13.0	
Linux 64-bit (x86-64)	Google Chrome 77.0 and later Mozilla Firefox 68.0 and later	Google Chrome 77.0 and later
Microsoft Windows 10 (64-bit)	Google Chrome 77.0 and later Mozilla Firefox 68.0 and later	Google Chrome 77.0 and later Mozilla Firefox 68.0 and later

Operating System	Web Browser (64-bit)	Web Browser (32-bit)
	Microsoft Edge on Chromium 80.0 and later	Microsoft Edge on Chromium 80.0 and later
Microsoft Windows 10 (32-bit)		Google Chrome 77.0 and later Mozilla Firefox 68.0 and later Microsoft Edge on Chromium 80.0 and later
Microsoft Windows 8.1 (64-bit)	Google Chrome 77.0 and later Mozilla Firefox 68.0 and later	Google Chrome 77.0 and later Mozilla Firefox 68.0 and later
Microsoft Windows 8.1 (32-bit)		Google Chrome 77.0 and later Mozilla Firefox 68.0 and later
Microsoft Windows 8 (64-bit)	Google Chrome 77.0 and later	Google Chrome 77.0 and later Mozilla Firefox 68.0 and later
Microsoft Windows 8 (32-bit)		Google Chrome 77.0 and later Mozilla Firefox 68.0 and later
Microsoft Windows 7 (64-bit)	Google Chrome 77.0 and later	Google Chrome 77.0 and later Mozilla Firefox 68.0 and later
Microsoft Windows 7 (32-bit)		Google Chrome 77.0 and later Mozilla Firefox 68.0 and later
Microsoft Windows Server 2012 R2		Google Chrome 77.0 and later Mozilla Firefox 68.0 and later
Microsoft Windows Server 2012		Google Chrome 77.0 and later Mozilla Firefox 68.0 and later
Microsoft Windows Server 2008 R2		Google Chrome 77.0 and later Mozilla Firefox 68.0 and later

Mobile Platform Support

Some SAS Viya user interfaces are not currently supported on mobile devices.

When SAS Visual Analytics Apps for iOS, Android, and Windows 10 are installed on a mobile device, users of the following interfaces can view and explore reports on the mobile device:

- SAS Visual Analytics

- SAS Visual Statistics
- SAS Visual Machine Learning

SAS Visual Analytics Apps are supported on iOS and Android smartphones and tablets, and on computers and tablets running Microsoft Windows 10. Each app is written specifically for its operating system and provides native support for these devices so that you can view and explore reports on your mobile device using a touchscreen. You can download the apps for free from the Apple App Store, Google Play, or the Microsoft Windows Store.

Screen Resolution

The minimum screen resolution for each client machine that will access the SAS Viya user interfaces is 1280 x 1024.

Support for Map Services

Some SAS Viya applications support multiple third-party map services, including OpenStreetMap, ArcGIS Online, and Esri ArcGIS Enterprise Portal. For Esri ArcGIS Enterprise Portal access, SAS Viya supports an Esri server version 10.4 and later. The server can be installed on premises (Esri ArcGIS Enterprise Portal) or running remotely (ArcGIS Online).

Esri has ended support for the 10.3.1 version as of December 2020. SAS might not be able to assist you if you encounter problems with versions earlier than 10.4.

IMPORTANT SAS Viya currently supports only token-based authentication for Esri. For example, an Esri server that is configured for Integrated Windows Authentication (IWA) is incompatible with SAS Viya.

Product-Specific Requirements

Limitations to Multi-Tenancy Support

Starting with 2021.1.6, SAS Viya can be deployed with multi-tenancy enabled. However, a few limitations apply to this feature at this time.

SAS Viya cannot be deployed with multi-tenancy enabled in a Red Hat OpenShift environment.

The following SAS product offerings do not support multi-tenant capabilities:

- SAS Analytics for IoT
- SAS Asset and Liability Management

- SAS Assortment Planning
- SAS Demand Planning
- SAS Event Stream Processing
- SAS Field Quality Analytics
- SAS Financial Management
- SAS Financial Planning
- SAS Inventory Optimization
- SAS Markdown Optimization
- SAS Production Quality Analytics
- SAS Revenue Optimization
- SAS Size Optimization

Requirements for SAS® for Microsoft® 365 Clients

SAS for Microsoft 365 enables SAS analytics to access reports directly from Microsoft Excel 365 and provides integrated features. Some requirements for SAS for Microsoft 365 differ from those of other SAS Viya offerings.

Supported Browsers for the Web Application

After you deploy SAS Viya and configure SAS for Microsoft 365, the SAS for Microsoft 365 web application is available from the Office 365 Excel web client in a supported web browser. When it runs in a browser, the SAS for Microsoft 365 web application supports only the following web browsers:

- Google Chrome 77.0 and later
- Microsoft Edge on Chromium 80.0 and later
- Mozilla Firefox 68.0 and later
- Apple Safari 13.0 and later on macOS

Browsers running on mobile or touchscreen devices are not supported at this time.

Desktop Application Requirements

SAS for Microsoft 365 can also run as an installed add-in to a desktop version of Microsoft Excel. In order to deploy the SAS for Microsoft 365 installed add-in for Microsoft Excel, use a machine that meets the following requirements:

- Microsoft Windows 10 or macOS
- Microsoft Excel: Microsoft 365 version 16.0.13530.20424 or later
- Microsoft Excel requires Microsoft Edge on Chromium with WebView2

  For more information, see this Microsoft article.

If you subscribe to the Semi-Annual Enterprise channel for Microsoft Office, an administrator must perform an additional step to enable the use of the WebView2 browser control in Office Version 2102

(July 2021). A new registry key is required. For details, see https://developer.microsoft.com/en-us/microsoft-365/blogs/understanding-office-add-ins-runtime/.

Security Requirements

Microsoft requires add-ins, such as SAS for Microsoft 365, to run in an iframe in the Office 365 web application. To ensure that SAS for Microsoft 365 works properly, your SAS administrator must update these properties in SAS Environment Manager after deployment. After changing properties in SAS Environment Manager, you must restart the SAS Viya pods for the changes to take effect.

For more information and step-by-step instructions, see "Configuring SAS for Microsoft 365" in *SAS for Microsoft 365: User's Guide*.

Requirements for SAS® Model Risk Management

SAS Model Risk Management only supports an internal PostgreSQL database for the SAS Infrastructure Data Server. For more information, see "PostgreSQL Deployment Options" on page 64.

SAS Model Risk Management supports extensive customization. A Git repository is therefore required to enable the management and versioning of changes and customizations that occur over the lifetime of the deployment. You can use a Git repository from GitLab or Microsoft GitHub.

Your Git repository must have at least one initial default branch (for example, `/main`) with at least one file (for example, Readme.md) in it.

The following information is required in order to set up the connection to the Git repository:

- the protocol that is used to connect to the Git repository (for example, https)
- the user ID that will be used to connect to the repository (for example, mrmadmin)
- the base64-encoded version of the personal access token for the Git user ID
- the name of the repository (for example, mrmrepository)
- the name of the branch (for example, main)
- whether you want to append the namespace name to the branch name (Y or N)
- the full URL to the Git repository, including the protocol (for example, `https://mygitrepo.company.com/mrmrepository.git`)
- the URL of the host of the Git repository (for example, mygitrepo.company.com)

Customizations are required to configure PVCs for SAS Model Risk Management. For more information, see "Specify PersistentVolumeClaims to Use ReadWriteMany StorageClass" in *SAS Viya: Deployment Guide*.

Requirements for SAS® Workload Management

If your deployment includes SAS Workload Management, the SAS Workload Orchestrator will be used by the launcher to schedule work that belongs to the compute workload class. To enable this capability, your cluster must include nodes that are labeled for that workload class. When they are used by other SAS offerings, compute servers are launched dynamically, on demand. However, when SAS Workload Management is deployed with SAS Viya, compute servers and other components that

are started by the launcher do not run if hosts with the `workload.sas.com/class=compute` label are not found in the cluster. For more information about workload classes, see "Plan the Workload Placement" in *SAS Viya: Deployment Guide*.

In addition, the SAS Workload Orchestrator ClusterRole/ClusterRoleBinding must be applied to the cluster. The ClusterRole/ClusterRoleBinding is used to get information about node resources, enabling SAS Workload Orchestrator to determine CPU and memory resources to be used for scheduling. After you have downloaded and uncompressed the deployment assets, you can find information about applying these settings in the README file at `$deploy/sas-bases/overlays/sas-workload-orchestrator/README.md`.

Verify the Environment

Run a Pre-installation Check

SAS recommends that you run the SAS Pre-Install Check Utility before deploying the software.

SAS Pre-Install Check retrieves information about your Kubernetes cluster and verifies that the cluster meets the documented requirements to deploy SAS Viya. The utility generates a report.html that displays the results of the checks that it performs.

The files that are required to create the Pre-Install Check Report and the instructions for running the utility are available at the SAS GitHub site. The README file contains links to "Pre-installation of SAS Viya System Requirements" and other tools that you can use once the deployment has completed.